

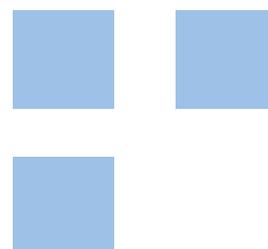


Department of Economics - FEA/USP

Predicting Skills of Runaway Slaves in São Paulo, 1854-1887

RENATO P. COLISTETE

WORKING PAPER SERIES N^o 2021-15



DEPARTMENT OF ECONOMICS, FEA-USP

WORKING PAPER N° 2021-15

Predicting Skills of Runaway Slaves in São Paulo, 1854-1887

Renato P. Colistete (rcolistete@usp.br)

Abstract

This paper examines the skills of enslaved labour during the second half of the nineteenth century in the province of São Paulo. The analysis is based on data from 3,376 individuals collected in advertisements of runaway slaves published by São Paulo newspapers between 1854 and 1887. As only a small part of the announcements listed runaways' occupations, we draw on individual details on sex, age, ethnicity, residence, physical characteristics and other features of fugitives with advertised occupations to predict the skills of the remaining subset of runaways, using classification algorithms from machine learning. Overall, both observed and predicted skilled runaways converged in their characteristics: skilled runaways were mostly male, older than their low-skilled counterparts and predominantly from farms and plantations, rather than urban settings. Africans were not at a disadvantage in artisanal jobs when compared with Brazilian-born runaways, and the skill gap between mixed-race and black fugitives was negligible. Although the enslaved population suffered from very low levels of literacy, the few runaways with an ability to read or write tended to work in more qualified and artisanal occupations, indicating that education may have been valuable even under the appalling conditions of slavery. These results are important both for the analysis of slavery in Brazil and comparisons with other plantation societies in the Americas.

Keywords: Slavery, Runaways, Skills, Skilled slaves, Machine learning, São Paulo, Brazil

JEL Codes: N3; N36; N5; N56; J24.

Predicting Skills of Runaway Slaves in São Paulo, 1854-1887

Renato P. Colistete
University of São Paulo
rcolistete@usp.br

15 April 2021

Abstract

This paper examines the skills of enslaved labour during the second half of the nineteenth century in the province of São Paulo. The analysis is based on data from 3,376 individuals collected in advertisements of runaway slaves published by São Paulo newspapers between 1854 and 1887. As only a small part of the announcements listed runaways' occupations, we draw on individual details on sex, age, ethnicity, residence, physical characteristics and other features of fugitives with advertised occupations to predict the skills of the remaining subset of runaways, using classification algorithms from machine learning. Overall, both observed and predicted skilled runaways converged in their characteristics: skilled runaways were mostly male, older than their low-skilled counterparts and predominantly from farms and plantations, rather than urban settings. Africans were not at a disadvantage in artisanal jobs when compared with Brazilian-born runaways, and the skill gap between mixed-race and black fugitives was negligible. Although the enslaved population suffered from very low levels of literacy, the few runaways with an ability to read or write tended to work in more qualified and artisanal occupations, indicating that education may have been valuable even under the appalling conditions of slavery. These results are important both for the analysis of slavery in Brazil and comparisons with other plantation societies in the Americas.

Introduction

In 1884, the US Vice-Consul in Rio de Janeiro visited the large Ibicaba plantation located in the centre of the fast-growing coffee region around Limeira, a municipality in the province of São Paulo. There, Christopher C. Andrews was introduced by Ibicaba's owner, José Vergueiro, to the plantation's modern buildings and equipment for coffee hulling, cleaning, polishing and other processing activities. "The machinery and works for cleaning and preparing [coffee]," wrote Andrews, "were of a character calculated to excite wonder and admiration." Besides the wonders of coffee technology, Andrews also noted a group of workers in charge of the equipment employed to process coffee after harvesting: a "number of his slaves are skilled mechanics, whom he has instructed."¹

The skilled slaves worked side by side with free Brazilians, Germans, Swiss and other European immigrants since the time Ibicaba plantation pioneered the introduction of European families in coffee plantations in the late 1840s, under a sharecropping (*parceria*) system that soon spread to other large farms in São Paulo. Although sharecropping collapsed after unrest among colonists that culminated in a high-profile rebellion by Swiss immigrants at Ibicaba in 1857, new types of contracts replaced *parceria* and European families continued to flow to work in coffee plantations filled with enslaved people until

¹ Andrews, *Brazil*, 153–54.

the abolition of slavery in 1888.² At the time of Andrews's visit in 1884, there were about 400 slaves and 50 immigrants in the daily work routine of Ibicaba, according to another visitor, Mauricio Lamberg.³ Unfortunately, we do not know yet how many enslaved workers were mechanics or employed in skilled trades in this plantation in 1884. In any event, skilled slaves were recorded on other large farms in the coffee provinces at about the same time of Andrews's visit. Louis Couty, for example, reported a median of 5 percent of craftsmen among healthy adult slaves in nine plantations on whom he was able to gather first-hand information from their owners.⁴

Scholars have long investigated the features, roles and numbers of skilled slaves in practically all plantation societies of the Americas, although with considerable variation across regions, products and time periods. A rich set of sources including probates, estate accounts, advertisements, local surveys and censuses has been employed to analyse the slave occupational structure of the US Antebellum South's cotton enterprises, among others.⁵ Research on sugar plantations in different parts of the Caribbean has also addressed the enslaved workforce and its hierarchies in great detail.⁶ Brazil's sugar plantations, on the other hand, have received relatively little attention in this respect, despite some noticeable exceptions.⁷ Perhaps paradoxically in view of the time period and importance, coffee plantations of nineteenth-century Brazil are even less well-known in terms of composition of labour force, occupations and skills.⁸ Urban slave skills have attracted more attention by historians, although only a fraction of major cities and regions of Brazil has been covered by more extensive studies.⁹

This article is an attempt to draw a more detailed and nuanced picture of skilled slaves who worked in both rural and urban areas in São Paulo, the province in Brazil which became the leading coffee export region in the world from the late nineteenth century. The analysis is based on a rich set of information on individuals collected from advertisements of runaway slaves published by the São Paulo press during the second half of the nineteenth century. The data on enslaved people include sex, age, birthplace, residence, physical details such as skin colour, height, body constitution as well as personal abilities of reading, writing and speech, along with labour skills and occupations. A total of 3,376 fugitive slaves were recorded in the announcements in newspapers of the provincial capital, São Paulo, between 1854 and 1887, the year before the final abolition of slavery in Brazil. As expected, however, the information available in advertisements was irregular and fragmented. For example, only 2,488 (73.7%) of runaways could be traced back in their birthplace, whether Brazil or Africa. The data from the advertisements are especially limited with respect to runaways' occupations and skills. Only 994 (29.4%) out of 3,376 fugitive slaves were recorded with information that could be used to identify their work activities and specialisations.

We take advantage of the richness of the data on individual runaway slaves in an attempt to extend our knowledge of their skills beyond the recorded information available in the advertisements. First,

² Holanda, "Prefácio"; Dean, *Rio Claro*, ch. 4; Witzel de Souza, "Bonded Laborers," ch. 2.

³ Lamberg, *O Brazil*, 327.

⁴ Calculated from Couty, *Étude de Biologie Industrielle*, 97.

⁵ Fogel and Engerman, *Time on the Cross*, 38–43; Menard, "Maryland Slave Population"; Kulikoff, *Tobacco and Slaves*, ch. 10; Johnson, "Work"; Marks, "Skilled Blacks"; Fogel, *Without Consent or Contract*, ch. 2; Linsin, "Skilled Slave Labor"; Dunn, *Tale of Two Plantations*; Pritchett and Hayes, "Occupations of Slaves."

⁶ For the British Caribbean, for example: Hall, "Slaves and Slavery"; Craton, *Searching*, chs 5 and 12; Ward, *British West Indian Slavery*; Sheridan, "Changing Sugar Technology"; Higman, *Slave Populations*, ch. 6; Dunn, *Sugar and Slaves*, chs 6-7; Roberts, "Working Between the Lines"; Mair, *Historical Study of Women*, ch. 6; Dunn, *Tale of Two Plantations*.

⁷ Graham, "Slave Families"; Schwartz, *Sugar Plantations*, chs 6, 12-4; Barickman, "Persistence and Decline"; Ferlini, *Terra, trabalho e poder*, ch. 3; Godoy, "Fazendas Diversificadas"; Schwartz, "A Commonwealth Within Itself"; Fraga, *Crossroads of Freedom*, ch. 1.

⁸ Although not dealing specifically with slave occupations and skills, see: Slenes, "Demography and Economics", pp. 530-42; Dean, *Rio Claro*, ch. 3; Florentino and Góes, "Parentesco"; Oliveira, "Forasteiros", ch. 1. For the slave routine, discipline and conditions in coffee plantations, see Crüwell, *Brazil as a Coffee-Growing Country*, ch. 3; Stein, *Vassouras*, ch. 5-7.

⁹ Lobo et al., "Estudo"; Dias, "Fímbricas"; Karasch, *Slave Life*, ch. 7; Andrade, *Mão de Obra Escrava*; Wissenbach, "Arranjos"; Nishida, "Manumission"; Lima, "Escravos Artesãos"; Lima, *Artífices*; Soares, *O "Povo de Cam"*; Read, *Hierarchies of Slavery*, ch. 4.

we present an overall picture of the skills of fugitive slaves. This analysis relies on the relatively small number of runaways whose occupations were informed in the announcements. Second, we draw on individual data on sex, age, ethnicity, residence, physical characteristics and other features to predict the skills of the remaining set of runaways, that is, the 2,382 enslaved persons who ran away from their owners and about whom there is no information regarding their jobs and skills. The prediction for the group with unknown skills is conducted with classification algorithms from machine learning, which are tested in their precision to label the slaves lacking job details by using the features already known about the runaways with occupational data. In the end, we will be able to figure out how many of the fugitive slaves without occupational information in the advertisements could have been skilled workers. Moreover, it will be possible to describe the main individual characteristics of the otherwise unknown, but labelled skilled runaways, comparing them with the features of the skilled slaves announced in the press.

The findings indicate that skilled runaways in São Paulo were overwhelmingly male, older than their low-skilled counterparts and predominantly from the rural areas of farms and plantations, rather than urban settings. Among other features related to the division of labour and specialisation, Africans seem to have not been disfavoured in the allocation of artisanal jobs, and perhaps they may even have had an edge over Brazilian-born runaways in skilled jobs. The skill gap between mixed-race and black fugitives was small and skin colour was probably not a significant determiner of skills. Despite literacy levels below 1 percent among the enslaved population in São Paulo, artisans were more literate than non-artisans and those few runaways who were able to read and write may have enjoyed an advantage in learning and working in more qualified occupations. As we will see, some of these results are at variance with what has been reported for other plantation economies and may have an importance beyond the specific case of the São Paulo province in the nineteenth century.

The next section presents the dataset of runaway advertisements, the procedures for collecting the data and the criteria adopted for defining skill categories. We then briefly analyse the features of skilled and low-skilled slaves as recorded in the newspapers. The sections then following deal with machine learning algorithms and the steps to classify the group of runaways with unknown skills, by using the individual features of those with advertised occupations. Then the group of enslaved workers predicted to be skilled by the best model is compared with the known set of skilled runaways. The final section summarises the findings and conclusions.

Advertisement data

The decision to run away represented one of the most dramatic forms of resistance which enslaved Africans and descendants could attempt in order to evade the harsh reality of slavery. Escaping alone or in groups from plantations, small farms or urban places was often a risky, dangerous and uncertain enterprise, which broke family ties, disrupted means of subsistence, and posed a real threat to the wellbeing and life of fugitive slaves. Despite all these conditions, flights were a frequent and daily part of the operation of plantations, other economic units and households. From the owners' point of view any runaway was a challenge to the discipline and control that they sought to impose on the enslaved labour force. Besides, fugitives implied unrealised production and a potential capital loss, which could only be avoided through extra costs in search and capture – and not always with success. Announcements of fugitive slaves in the local press were thus one of the ways by which owners attempted to assert their challenged authority and mitigate economic losses.¹⁰

A vast literature has assessed the advantages and limitations of advertisements as a source for studying

¹⁰ Conrad, *Últimos Anos*, 18–22; Johnson, “Runaway Slaves,” 418; Heuman, “Runaway Slaves,” 95; Gebara, “Escravidão,” 127–28, 144; Schweninger, “Counting the Costs”; Franklin and Schweninger, *Runaway Slaves*, chs 1–3, 9, 11; Sidbury, “Slave Resistance”; Dittmar and Naidu, “Contested Property”; Foner, *Gateway to Freedom*, 4–6.

multiple aspects of slavery.¹¹ The problems are numerous. Flights announced in newspapers possibly referred only to a tiny part of the total number of runaways. Most fugitive slaves tended to be male, leaving women heavily underrepresented in the sample. Furthermore, newspapers were uneven in their geographical coverage, what becomes a drawback in societies with low literacy and diffusion of the press. The costs of announcements, often published over consecutive days to be more effective, introduced a bias towards wealthy masters and more valuable slaves. A high chance of capturing fugitives in the neighbourhood areas with the help of policemen or slave catchers made advertisements a last resort, perhaps even an exceptional alternative. Nonetheless, the advantages of runaway advertisements as a historical source are also significant. These announcements described characteristics of enslaved people uncommon in other sources, such as physical details, cloths, abilities to speak, accents, diseases, injuries, scars, literacy and work specialisations. Owners aimed to retrieve fugitives and sought to be objective in the intelligence given to newspapers' readers, many times describing traits as special as the slaves' ways of walking and their gifts as singers, instrument players or dancers. It is therefore not surprising that advertisements of runaways were from an early date recognised as a valuable source for studying different aspects of slavery. In Brazil, for example, after a first article on the subject in 1935, Gilberto Freyre published in 1963 a detailed study on advertisements of runaways in provincial newspapers during the nineteenth century.¹²

The primary database used in this article is composed of advertisements of runaway slaves that appeared in four newspapers published in São Paulo during the second half of the nineteenth century: *Correio Paulistano* (1854-1887), *Diario de S. Paulo* (1865-1878), *Jornal da Tarde* (1878-1881) and *Provincia de São Paulo* (1875-1887).¹³ All these newspapers were based in the capital, São Paulo, but aimed at reaching a readership across the whole province. Every available issue of the newspapers was examined, although there were two missing years – 1860 and 1861 – in the collection which could not be recovered. Enslaved persons who ran away alone or in a group were coded as a single case with his or her individual characteristics and other related details, such as the owner's name and place of residence. The data cleaning and preparation were done with the pandas library in the Python programming language.¹⁴ Overall, a total of 3,276 individual fugitive slaves were retrieved from the newspapers, distributed between 1854 and 1887.

A series of steps were taken to avoid duplication, to fill gaps in missing data and to reduce errors in the preparation of the dataset. A difficult issue was dealing with announcements running for several days or even over alternate months, posing the need to distinguish such cases from truly recurrent fugitives. As enslaved persons usually did not have surnames and were only known by common Brazilian first names, the search for repeated cases had to be linked to personal traits, the name of owners and other individual details. With regards to those who did escape more than once, the criterion was to keep only the first flight to avoid double counting of the same person as a runaway. Several features had to be adjusted or transformed before analysis. For example, the slaves' ages were usually reported in a rather imprecise way in newspapers, sometimes with intervals as vague as "between 18 and 22 years." In such cases, a simple average was calculated as an approximation. When just an ambiguous reference was made ("more or less 25 years old" for example), the specific age cited (25 years in this case) was taken as a proxy for the real age of the runaway slave.

In other situations the adjustments required more complex procedures or conventions. Slave's heights appeared in the advertisements under generic labels such as "tall" and "regular" as well as with

¹¹ White, "Black Fugitives," 25-27; Heuman, "Runaway Slaves," 97-100; Prude, "Runaway Ads," 124-26; Waldstreicher, "Reading the Runaways," 243-49; Franklin and Schweninger, *Runaway Slaves*, 170-78; Wallace, "Fugitive Slave Advertisements."

¹² Freyre, *Escravo nos Anúncios*.

¹³ These newspapers are available online in the Brazilian National Library' digital collection (*Hemeroteca Digital*, <https://bdigital.bn.gov.br/hemeroteca-digital>) and *O Estado de São Paulo* portal (<https://acervo.estadao.com.br>).

¹⁴ McKinney, "Data Structures"; Van Rossum and Drake, *Python* 3.

numbers in centimetres, inches and palms. All numeric heights were first converted to centimetres and then classified as “short,” “regular” and “tall,” along with those so defined by announcers. This classification was based on the anthropometric studies by Daniel Franken and Joerg Baten *et al.*, who found that the average stature of Brazilians lay somewhere between 1.64 and 1.66 m during the period 1810-1880.¹⁵ Heights between these numbers were regarded as “regular” and those below and above that interval were considered “short” and “tall,” respectively. Skin colour also demanded simplification and conventions. The search of newspaper announcements turned up more than 20 unique descriptors for skin colours of fugitive slaves, with mixed-race and gradations such as “white” (“branca”), “almost white” (“quase branca”), “light brown” (“parda clara”) and “dark brown” (“parda escura”). As there was no common pattern in the advertisements to define mixed-races, the skin colours were classified in two simple categories: mixed-race (“parda”) and black (“preta”).¹⁶ The body constitution of fugitive slaves was another feature with dozens of specific terms, which were often ambiguous in their meaning. All of them were codified according to four labels: slim (e.g., “magro,” “franzino”), regular (“regular,” “bem feito”), stout (“cheio de corpo,” “grosso,” “corpulento”) and robust (“entroncado,” “musculoso”).¹⁷

Another set of characteristics was relatively straightforward to interpret and codify. The announcements frequently informed whether runaways were born in Africa (“de nação”) or Brazil (“crioulos”). This is the primary ethnic division adopted in the following analysis.¹⁸ Brazilian-born slaves were also known by their province of origin or by generic reference to the areas from which they came in Brazil – “south” and “north” regions for example. Usually the advertisements gave the name of the municipality where the runaways had lived and worked, although other related details were sometimes missing. When there was enough information about where they resided – for example in plantations, small farms or urban homes – runaways were also classified according to their dwellings in urban or rural areas. Literacy was another major piece of information which could be gleaned more easily from advertisements. Every fugitive slave identified as being able to read and/or write was deemed as having achieved a level of literacy which qualified him or her as literate. When there was no reference to literacy, it was assumed that the runaway was illiterate.

Despite the richness of the data, a further problem with the advertisements is that the information provided for each feature of runaway slaves was imbalanced. While name, gender and municipality could be identified for all runaways, traits such as skin colour (90.1%), height (86.3%) and age (81.1%) were only available for a part, although substantial, of the individuals. For some characteristics the information was scarcer: whether runaways resided in urban or rural areas could be recovered for 61.9 percent of the cases, whereas body constitution (57.2%) and rewards (43.2%) resulted in even smaller figures. For occupations, as already mentioned, the information was much more sparse, with only 29.4 percent of all fugitive slaves having some evidence for the jobs in which they laboured on plantations and farms, and in shops and homes. The uneven number of observations for each feature of the dataset is shown in the Appendix, Figure A1.

The data on occupations provide the evidence needed to label the runaways according to their skills, even though the classification of skills in a rural and slave-based society is a rather loose task. First,

¹⁵ Baten, Pelger, and Twrdek, “Anthropometric History”, table 2, p. 321, used male prisoners data from Rio de Janeiro; Franken, “Anthropometric History”, p. 390, was based on male soldiers of the Brazilian army recruited in Rio de Janeiro. Although these figures are from free male soldiers or convicts, it is likely that the notions of stature entertained by announcers were similar regarding free and enslaved people.

¹⁶ On the concepts of mixed-race and skin colour in Brazil, see Lowrie, “Negro Element”; Degler, *Neither Black Nor White*; Piza and Rosemberg, “Cor nos Censos”.

¹⁷ The terms used for the classification of body constitution relied on contemporary dictionaries: Silva Pinto, *Diccionario*; Caldas Aulete, *Diccionario Contemporaneo*.

¹⁸ Africans were further split into quite different cultural, linguistic, religious, and political groups. References to African nations and ethnic divisions in the advertisements were common although very often imprecise, with generic terms like “de nação,” “African,” “Guiné” and “da Costa.” See Karasch, *Slave Life*, pp. 8-28, 304-6; Reis, “Identidade”; Slenes, ““Malungu, Ngoma Vem!””; Mattos, “De Cassange, Mina, Benguela a Gentio da Guiné”, ch. 2; Gomes, “Demografia Atlântica”.

enslaved workers usually engaged in a range of activities either in their regular working day, cropping seasons or their life cycle. Even those with more specialised functions could split their work schedule or be allocated to different jobs throughout the year.¹⁹ This variety of work experience was clearly reflected in the advertisements of fugitive slaves, with several of them being recorded as carrying out multiple occupations. Second, the overall definition of slave (and free labour) skills relies on the knowledge and abilities which are viewed as necessary for the execution of specific tasks in plantations, shops and the like. In practical terms, the concept of skilled trades adopted by historians has varied from a simple separation between agricultural (non-skilled) and non-agricultural (skilled) jobs to more elaborate notions of non-skilled, low-skilled and high-skilled occupations in view of the training, experience and dexterity required.²⁰

For this dataset, a straightforward definition was adopted based on the notion of artisanal work: runaway slaves were considered skilled when working in artisanal and craft occupations in both rural and urban areas, while the remaining slaves were entered as low skilled. Typical artisan slaves included blacksmiths, masons, mechanics and carpenters, who usually had to be trained on the job for a substantial length of their lifetimes and were acknowledged for their abilities. Runaways were assigned as skilled workers both when the advertisements mentioned just one artisanal occupation, and when besides this craft trade the description included low-skilled jobs they were able to perform, such as agricultural labour and dealing with animals.²¹ The skilled runaways made up a small fraction (462 or 13.7%) of all runaway slaves in newspaper advertisements between 1854 and 1887, as shown in Figure 1.

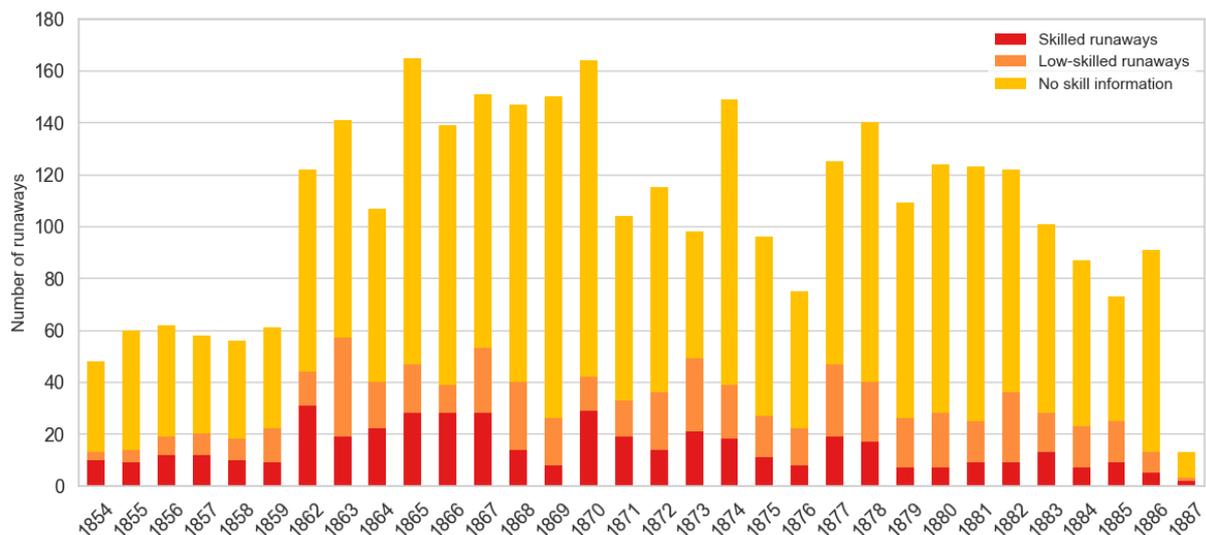


Figure 1: Skills of runaway slaves in newspaper advertisements, province of São Paulo, 1854-1887
 Source: advertisements published in the *Correio Paulistano*, *Diario de S. Paulo*, *Jornal da Tarde* and *Provincia de São Paulo*.

On average, skilled slaves accounted for 46.5 percent of all fugitives with informed occupations, a share that is well above what other sources usually recorded as skilled workers among the enslaved population. Adopting a similar definition of skills based on artisanal trades, the national census of

¹⁹ Fallope, “Les Occupations d’Esclaves,” 194–95; Higman, *Slave Populations*, 164, 170–71; Linsin, “Skilled Slave Labor,” 187–88; Roberts, “Working Between the Lines,” 554, 572–75; Penna, “Escravidão,” 83–101; Mattos, *Das Cores do Silêncio*, 57–60.

²⁰ Berlin and Gutman, “Natives and Immigrants,” 7; Johnson, “Work,” 331–34; Marks, “Skilled Blacks,” 540–44; Irwin, “Farmers and Laborers,” 55–56; Higman, *Slave Populations*, 170–72; Hardesty, “‘The Negro at the Gate’,” 78–88; Florentino and Góes, “Parentesco,” 194–95; Lima, “Escravos Artesãos,” 474.

²¹ Table A1 in the Appendix shows a summary of the occupations – skilled and low skilled – found in the advertisements and logged in the dataset.

1872 exhibited only 4.3 percent of skilled workers among the bonded population over the age of ten in the province of São Paulo. The share of skilled artisans in the free population over ten years was almost the same, 4.7 percent.²² The 1872 census may have singled out mainly urban craftsmen and left unrecorded most artisans working in rural establishments, as it is plausible, but the difference between the percentages (4.3% against 46.5% skilled runaways) is too high to be only due to a lack of coverage or a conceptual limitation.²³ Another source, the special slave *matrículas* of 1873, followed a different and even more vague definition of occupations than the national census and found that skilled slaves made up 15.0 percent of the population over thirteen years in the province of São Paulo at the time.²⁴ These are indications that skilled slaves were disproportionately represented in the advertisements for runaways, possibly because of their valuable abilities and importance for the work process.²⁵ The bias towards skilled runaways therefore should be taken into account in the following comparisons with the group of low-skilled slaves.

Who were the skilled runaways?

The skilled runaways announced in newspapers were concentrated in a few regions and municipalities of São Paulo province, a geographical distribution nearly symmetrical to that of the whole group of fugitive slaves. Figure 2 shows the number of artisan runaways in municipalities according to São Paulo's administrative division in 1872. In relative terms, the Capital had the largest share (25.1%) of skilled runaways advertised among the province's municipalities, which reflects the bias of the newspapers located in the provincial capital. This share is much larger than the proportion of enslaved population over the age of ten years living in the city of São Paulo – 3.7 percent in 1872, for example. Other localities with a substantial number of skilled fugitives were situated in regions intensive in coffee production, both old and new centres which grew fast in the last decades of slavery: Campinas (15.2%), Itatiba (4.2%), Piracicaba (4.3%), Itu (3.9%), Jundiaí (3.7%), Mogi Mirim (3.5%), Limeira (3.0%) and Rio Claro (2.8%).²⁶

Most enslaved persons escaped from plantations and farms (72.3% of all fugitives) and it is worth noticing that the majority (69.0%) of the skilled runaways also came from rural areas, contrary to the notion that artisanal trades were essentially an urban activity. In fact, the operation of a coffee enterprise demanded a variety of crafts skills such as woodworking, construction and mechanics.²⁷ But when looked at according to their domicile, runaways from urban settings tended to be more skilled (50.3%) than those absconding from rural areas (42.5%). In the same fashion, skilled runaways were more likely to come from urban nuclei (31.0%) than rural areas (69.0%) when compared with the low-skilled fugitives (24.8 and 75.2%, respectively).

As has been noted, most runaways were males, which made women strongly underrepresented in advertisements. In São Paulo newspapers, only 11.0 percent of fugitives were females.²⁸ In addition,

²² Calculated from Brazil, *Recenseamento Geral de 1872*.

²³ On the classification of skills and rural population in the 1872 Census, Dean, *Rio Claro*, ch. 3; Castro, *Sul da História*, 96-101; Rodarte and Júnior, "Estrutura Ocupacional"; Lago, *Da Escravidão ao Trabalho Livre*.

²⁴ Brazil, *Relatório e Trabalhos Estatísticos de 1875*.

²⁵ For the bias of advertisements in favour of skilled runaways: Heuman, "Runaway Slaves", 97-9; Read and Zimmerman, "Freedom for Too Few", 406; Wallace, "Fugitive Slave Advertisements", 151-4.

²⁶ These municipalities comprised the three intervals with the highest number of runaways in Figure 2. The five classes in which the sample is split in the map were obtained with the Fisher Jenks Natural Breaks algorithm, which minimises the deviations from the mean within each range, while maximising the deviations from the means of the other classes. The estimation was done with the PySAL module (Rey and Anselin, "PySAL") and the map created with the geopandas (Jordahl et al., "Geopandas") and matplotlib libraries (Hunter, "Matplotlib").

²⁷ Couty, *Étude de Biologie Industrielle*, 96-101; Laërne, *Brazil and Java*, 310-21; Stein, *Vassouras*, 165, 233-36; Fragoso, "Sistemas Agrários," 61-72; Oliveira, "Forasteiros," ch. 1.

²⁸ For similar cases, White, "Black Fugitives", 26; Johnson, "Runaway Slaves", 424; Heuman, "Runaway Slaves", 98; Franklin and Schwening, *Runaway Slaves*, ch. 9; Gomes, "Jogando a Rede", 9-10; Amantino, "Escravos Fugitivos", 61-2; Machado, "Com

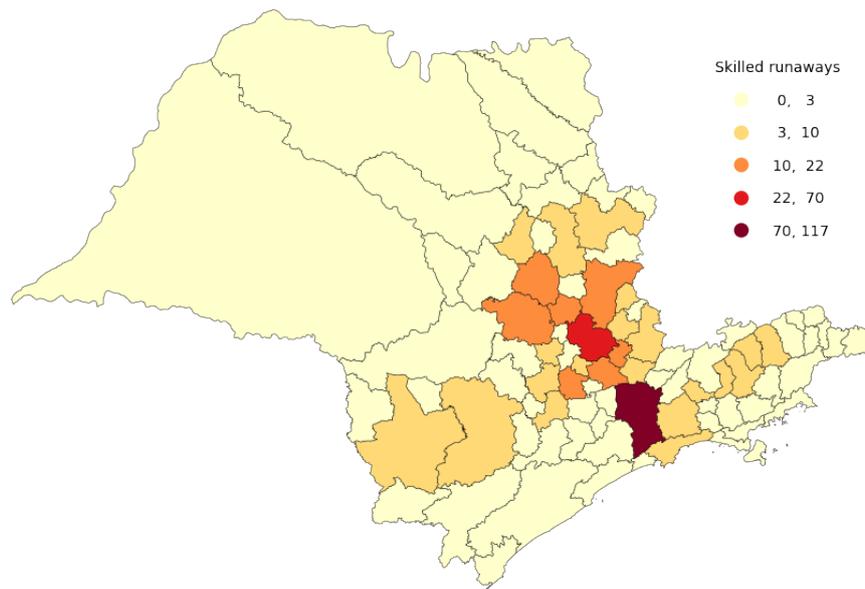


Figure 2: Geographical distribution of skilled runaway slaves, province of São Paulo, 1854-1887
Source: advertisements published in São Paulo newspapers.

artisanal trades have been documented as nearly exclusively male occupations in several plantation societies – and São Paulo was no exception.²⁹ Within the group with advertised occupations in São Paulo, the share of females was even smaller (3.7%) than in the sample as a whole, with only one woman (0.2%) being identified as working in an artisanal occupation.

Ethnicity was another aspect with a possible effect on skills in plantation economies.³⁰ About 88.9 percent of all runaways were born in Brazil (creoles), a figure mirrored in the high share of creole craftsmen (89.5%) encountered among fugitives with occupational data. However, when ethnic groups are considered separately, African (“de nação”) runaways had only a slightly lower proportion (44.7%) of artisans than Brazilian-born slaves (47.4%). This is somewhat against the likelihood that has been reported or assumed that African enslaved workers faced restrictions such as language barriers and were prejudiced in favour of creoles, becoming thus less represented in skilled occupations than Brazilian-born slaves – for example in nineteenth-century Campinas and Minas Gerais.³¹ A chi-square test of independence between samples shows that there is no statistically significant association between skills of creole and African runaways. This result suggests that Brazilian-born runaways may have been slightly more skilled because of conditions other than any preference against Africans.³²

Vistas à Liberdade”, 103-4; Costa, “Fugas de Escravos”, 92-3.

²⁹ Fogel, *Without Consent or Contract*, 42; Higman, *Slave Populations*, 189, 192; Mair, *Historical Study of Women*, 198–210; Cowling, *Conceiving Freedom*, 31–32; Santos, “Mothering Slaves,” 8.

³⁰ White, “Black Fugitives,” 29; Ward, *British West Indian Slavery*, 226–28; Fogel, *Without Consent or Contract*, 47–49; Higman, *Slave Populations*, 197–99.

³¹ Slenes, “Brazilian Internal Slave Trade,” 355; Bergad, *Slavery*, 185.

³² As the calculated chi-square statistic is smaller (0.141) than the critical point (3.841) at 5 percent significance level with one degree of freedom in the chi-square probability table, the test fails to reject the null hypothesis that there is no statistically significant association between the samples. The following references to the chi-square test follows the same criteria and will quote only the test statistic. For the use of the chi-square test and contingency tables, Siegel, *Nonparametric Statistics*, 175-9; Feinstein and Thomas, *Making History Count*, 202-10. The tests were performed with the SciPy module, Virtanen et al., “SciPy 1.0”.

A regional perspective reinforces the possibility that Africans were not at a disadvantage in artisanal occupations. Whereas in the Capital 59.1 percent of creoles mastered artisanal skills in comparison with only 32.4 percent of Africans, in a new and fast-growing coffee region such as Mogiana about 66.7 percent of African runaways were skilled, a much higher figure than the share of artisan slaves among creole runaways (44.3%). The same profile was observed in another area of expanding coffee production, the Central region, where skilled slaves made up 61.9 percent of all African runaways compared with 51.0 percent for creoles. Interestingly, the old coffee region of the Paraíba Valley exhibited a distribution more alike that of the Capital: skilled artisans accounted for 39.3 percent of creole runaways and just 28.6 percent of African runaways.³³

The skin colour of runaways appears to have a more distinctive pattern in terms of skills. Mixed-race (*pardos*) and black (*pretos*) runaways had similar proportions in São Paulo – 50.4 and 49.6 percent, respectively – but relatively more *pardos* (49.5%) were skilled artisans than *pretos* (44.4%). This difference meant that *pretos* formed the majority of low-skilled runaways (51.7%) and a smaller fraction (46.6%) of skilled runaways. Despite that, such a skill gap between *pardos* and *pretos* seems to be relatively narrow in comparison to similar slave economies, even though the evidence is dubious for other plantation societies in the Americas.³⁴ In the same way as ethnicity, the relationship between the skills of mixed-race and black runaways is not statistically significant, implying that artisanal jobs may not have been allocated on the basis of slaves' skin colour.³⁵ Once again, the results are quite heterogeneous from a regional perspective, but this time not between the Capital region and the coffee areas as happened in the case of ethnicity. In the Capital, 55.0 percent of mixed-race runaways were skilled compared to 47.5 percent of blacks. The discrepancy was even bigger in the Mogiana coffee region – 52.2 and 31.9 percent for mixed race and blacks, respectively. The region that most diverged from the overall trend was the oldest area of coffee plantations in São Paulo, the Paraíba Valley, where *preto* runaways (44.7%) were even more skilled than *pardos* (44.4%).

As expected, skilled runaways were significantly older than their low-skilled counterparts, among other possible reasons because of the experience and training required for artisanal trades.³⁶ The average age of craftsmen was 31 years and the median 30, whereas low-skilled runaways were 29 and 27 years of mean and median ages, respectively. Likewise, 27.8 percent of the skilled fugitives were in the age group from 35 to 54 years, well above the proportion for low-skilled runaways (22.0%).³⁷ The coffee zone (Central) with the highest relative share of skilled fugitives showed a mean age of 31 years (median 30), against 28 years (median 28) by low-skilled runaways – figures similar to those of other regions of intensive coffee production in the province of São Paulo. The Capital region, however, once more showed an outcome at odds with the coffee areas: its average age of skilled slaves (30) was below and not statistically different from that of low-skilled runaways (31).

Physical characteristics of height and body constitution had a potential influence on the distribution of skills.³⁸ Yet São Paulo runaways with regular stature had just a marginally higher proportion (50.2%) of skilled jobs than tall slaves (49.2%). Despite the fact that the gap was larger in relation to short runaway slaves (44.6%), the difference is not statistically significant. For body features the outcome is

³³ The regional division of São Paulo follows the original classification proposed by Milliet, *Roteiro do Café*. See Colistete, "Regiões e Especialização", 333-5.

³⁴ Fogel, *Without Consent or Contract*, 48-49; Higman, *Slave Populations*, 189, 194-97.

³⁵ Chi-square statistic of 2.137, so that it does not reject the null hypothesis of no association between the samples.

³⁶ Fogel, *Without Consent or Contract*, 56-58; Johnson, "Work," 339; Marks, "Skilled Blacks," 552; Graham, "Slave Families," 392-93.

³⁷ The distributions of skilled and low-skilled ages do not follow the condition of normality. The Mann-Whitney and the Kolmogorov-Smirnov non-parametric tests for mean ranks and the whole distribution, respectively, indicate that both series are statistically different. From now on, any reference to statistical significance of continuous variables will refer to these non-parametric tests, unless stated otherwise. See Siegel, *Nonparametric Statistics*, 116-36. Tests performed with the SciPy module, Virtanen et al., "Scipy 1.0".

³⁸ Fogel, *Without Consent or Contract*, 47-48.

similar. Craftsmen among stout (*cheios de corpo*) slaves were relatively fewer (41.6%) than those found among slim (48.5%), regular (51.1%) and robust (50.0%) runaways, but the differences are not statistically significant, in the sense that such characteristics appear not to have influenced the definition of who learned and worked in artisanal jobs.³⁹

As for literacy, very few runaways – only 2.7 percent of the total – were announced in newspapers as being able to read and/or write.⁴⁰ This figure is bigger than what was recorded by the national census of 1872, which found less than 1 percent (0.07%) of literate persons among the entire enslaved population in the province of São Paulo. As a matter of fact, the latter figure is not reliable either, as the census registered not a single enslaved person able to read or write in major municipalities, such as the provincial capital, indicating that slave's literacy, although extremely low, may have been deliberately omitted or unrecorded by masters, census officials or slaves.

Literacy among runaways with advertised occupations reached 5.7%, with artisans substantially more literate (8.0%) than low-skilled (3.8%) fugitives. The chi-square test shows that there is a statistically significant association between fugitives' literacy and skills, which means that those runaways able to read and/or write may have been preferred for skilled trades.⁴¹ Education seems to have mattered even under the appalling conditions faced by bonded people in a rural and slave society. Regionally, the traditional coffee region of Paraíba Valley held the highest proportion (12.1%) of literate runaways: at least nearly two times more than the other coffee regions of Paulista (6.6%), Mogiana (4.1%) and Central (3.8%). The Capital zone (4.1%) showed also a relatively small share of runaways with an ability to read and/or write.

In any case, the relatively high rates of literacy among skilled runaways was far away even from the reality of the group of runaways as a whole announced in newspapers. The same is true for the enslaved population counted in the national census, whose literacy was only of 0.7 percent in the province of São Paulo in 1872. From a human capital perspective, slaves in São Paulo were, as noticed before, at a similar position as free workers in terms of skills, with both groups exhibiting almost identical shares (4.3 and 4.7%, respectively) of skilled workers. The abysmal difference in human capital lay in literacy, an outcome of owners' attitudes and public policy.

The features of skilled runaways summarised so far can now be put together, in an attempt to draw an approximate picture of the skills of enslaved workers whose jobs were omitted in the advertisements.

Predicting skills

The prediction of skills for runaways with unknown occupations is essentially a binary classification task, that is, one that will assign 0 for low-skilled slaves and 1 for skilled slaves. Using the Python module scikit-learn, we compare the outcome of a range of classification algorithms trained over the data for the smaller group of 994 fugitive slaves and applied to the larger set of 2,382 runaways lacking skill information.⁴² Gaussian Naive Bayes is a probability classifier based on the Bayes theorem that assumes a normal distribution for continuous data and independence of all features. Giving its simplified assumptions, the Gaussian Naive Bayes classifier can be seen as a baseline model for comparison. Logistic Regression is a binary classifier based on a sigmoid function that also gives a probabilistic interpretation of how much an observation fits in the classes predicted by the model. Decision Tree

³⁹ The outcomes remain statistically non-significant when the short and stout runaways are compared with other height and body types classified as a single group.

⁴⁰ For slave literacy, see Cornelius, *When I Can Read*; Bly, "'Pretends He Can Read'"; Karasch, *Slave Life*, pp. 214-9; Barros, "Negrinhos que por ahi Andão"; Graham, "Writing from the Margins"; Morais, "Ler e Escrever" Machado, "Com Vistas à Liberdade".

⁴¹ The chi-square statistic was 7.317; thus the test rejected the null hypothesis of independence between the samples.

⁴² Pedregosa et al., "Scikit-Learn."

Classifier consists of a non-parametric algorithm in which each feature is split, according to a predefined threshold, in binary trees sequentially until a label is assigned to the specific observation. Support Vector Classifier and its linear version, Linear Support Vector Classifier, aim to draw a line (hyperplane) with a proper margin that divides the binary classes in such a way that new data can be assigned to the right label. Finally, the Gradient Boosting Classifier consists of adding simple models (such as decision trees) recursively, one after the other, and adjusting each stage with information gathered from the previous model until the best classification is obtained. In this exercise, we used an improved implementation of Gradient Boosting – the XGBoost algorithm.⁴³

The data of the 994 runaways on which the models were trained comprise both categorical and numeric variables, most of them containing missing observations – see Appendix, Figure A2. Since the classifiers implemented by scikit-learn do not support missing and non-numeric values, the preparation stage of the database required the adoption of imputation techniques. Additionally, the imputed values were transformed into numeric values when necessary and then scaled before being used for prediction.⁴⁴

Among the categorical features, provinces from where runaways originated could be identified in only 28.7 percent of cases. As for numerical features, the proportion of runaways with monetary rewards was also quite low (48.0%). Both features were discarded in order to minimise the potential distortion introduced by imputed values in the underlying distribution of real data. The remaining runaways' features have at least 60 percent of observations recorded in the advertisements: region, municipality, area, gender, age, height, body constitution, colour, ethnicity and literacy (Figure A2, Appendix). Four other variables were also considered to complement the features of runaways directly available in the newspapers: the regional division of the province of São Paulo, which has already been employed above; the distance in kilometres from the provincial capital, São Paulo; the altitude of each locality, and the existence of a railway station in a municipality at the year of the flights. All these variables aim to capture the existence of productive specialisation, population clusters and other characteristics in the wider area where the runaways lived.

A final step in the definition of variables was to assess whether all the available features were relevant in predicting runaways' skills. We employed a non-parametric method – mutual information – from theory of information that measures the degree of dependence between two variables, that is, between each feature listed earlier and the targeted runaways' skills. An advantage of mutual information is that it is able to capture both linear and non-linear dependency between variables.⁴⁵ The mutual information method implemented by scikit-learn indicated that three features had zero information relevant for the classification of runaways' skills: body structure, distance from the provincial capital and presence of railway stations in localities. After dropping these irrelevant variables from the classification task, the remaining ten features were selected for training by the algorithms: region, municipality, area, gender, age, height, colour, ethnicity, literacy and altitude.

Descriptive statistics of both numerical and categorical variables are presented in Tables 1 and 2. The age of runaways, for example, ranged from 10 to 61 years, with a mean of 30 years and median of 28 indicating an asymmetric distribution, skewed to the right. When compared to the mean, the standard deviation (10 years) confirms the large variability in the age of runaways (Table 1).

With regards the categorical features, only region, municipality, gender and literacy contain the full

⁴³ Hastie, Tibshirani, and Friedman, *Elements of Statistical Learning*; Efron and Hastie, *Computer Age Statistical Inference*; Chen and Guestrin, "XGBoost."

⁴⁴ Missing values of the numerical features were imputed with the multivariate feature method (IterativeImputer class) provided by scikit-learn, while the categorical features were imputed using the most frequent value in each column. The resulting categorical features were transformed into numeric values with the Target Encoder method of the Category Encoders package (McGinnis et al., "Category Encoders"). Finally, all features were scaled using scikit-learn's PowerTransformer with the Yeo-Johnson method to minimise variance and skewness of the data.

⁴⁵ Kuhn and Johnson, *Applied Predictive Modeling*, 476–77.

Table 1: Descriptive statistics of numerical variables, runaway slaves, São Paulo, 1854-1887

Measure	Age	Altitude
Count	788	994
Mean	30	681
Standard Deviation	10	114
Minimum	10	5
25%	23	591
50%	28	689
75%	35	784
Maximum	61	996

Source: advertisements in São Paulo newspapers.

set of observations (count equals 994). For instance, there were runaways from 67 different localities, with the capital, São Paulo, being the top municipality in the province from which enslaved individuals with recorded occupations (235 of them) ran away, as shown in the advertisements. The Central zone in turn was the region with the highest number (340) of runaways with occupations in the sample. Basic statistics of other categorical features can be seen in Table 2.

Table 2: Descriptive statistics of categorical variables, runaway slaves, São Paulo, 1854-1887

Measure	Region	Municipality	Area	Gender	Height	Color	Country	Literacy
Count	994	994	620	994	885	901	771	994
Unique	10	67	2	2	3	2	2	2
Top	Central	São Paulo	rural	male	regular	mixed	Brazil	no
Top frequency	340	235	449	957	420	457	686	937

Source: advertisements in São Paulo newspapers.

We can turn now to the classification performance according to the models trained with the runaways dataset. The group of 994 fugitives with occupations was split into a training set, which was used to fit a model, and a testing set, to which this model was applied and its predictions compared with the observed values. Each model was obtained after automatic optimisation with grid or random search in order to find the best set of parameters. The results are summarised in Table 3.

Table 3: Evaluation metrics of alternative models, runaway slaves, São Paulo, 1854-1887

Model	Accuracy	Precision	Recall	F1	Matthews coeff.
Gaussian Naive Bayes	0.65	0.69	0.65	0.65	0.35
Logistic Regression Classifier	0.67	0.68	0.67	0.67	0.35
Decision Tree Classifier	0.62	0.62	0.62	0.62	0.23
Linear Support Vector Classifier	0.68	0.68	0.68	0.68	0.36
Support Vector Classifier	0.66	0.66	0.66	0.65	0.30
XGBoost Classifier	0.71	0.71	0.71	0.72	0.44

Note: sample of 994 runaways split into training size = 75% and test size = 25%.

All the evaluation metrics converge to the XGBoost classifier as the model with the best performance in predicting the skills of runaways in the test set. These metrics refer to true and false predictions of positive (defined as skilled) and negative (low skilled) observations. Accuracy is the fraction of all correct predictions (true positives and true negatives) in relation to the total number of samples.

Precision gives the proportion of true positives predicted that are really positive. Recall shows how many true positives are predicted correctly but penalises false negatives. The f1-score tries to find a balance by calculating the harmonic mean between precision and recall. Finally, the Matthews coefficient measures the correlation between predicted and observed samples and is usually regarded as a metric that provides a more balanced account of all true and false positives and negatives. Whereas accuracy, precision, recall and f1 scores vary between 0 and 1, Matthews coefficient's values range from -1 to +1: a score of +1 indicates perfect prediction, 0 implies a random guess and -1 represents perfect misclassification between predicted and actual values.⁴⁶ As shown in Table 3, the scores of the XGBoost classifier are substantially higher than those returned by the second best models, such as the Linear Support Vector and Logistic Regression classifiers. The worst performance achieved was by the Decision Tree classifier, for instance when seen through the Matthews coefficient (0.23 against 0.44).

Figure 3 brings together confusion matrices that help to evaluate the binary classification of low-skilled and skilled runaways predicted by the best model, the XGBoost classifier, and the other algorithms. The rows of a 2x2 matrix correspond to the true classes, while the columns represent the predicted classes. The diagonal of the matrix shows the correctly classified elements.⁴⁷ As we have defined skilled as the positive class and low skilled as the negative class, and considering the XGBoost classifier as an illustration, the top left counts those runaways (68%) correctly classified by the model as low skilled (true negatives) and the bottom right measures the proportion of those correctly predicted (76%) as skilled (true positives). The other two entries of the matrix signify the number of elements that were wrongly classified as belonging to another class. Thus 24 percent of the runaways predicted as low skilled at the left bottom were in fact skilled (false negatives), whereas 32 percent of those classified as skilled at the top right turned out to be low skilled (false positives).

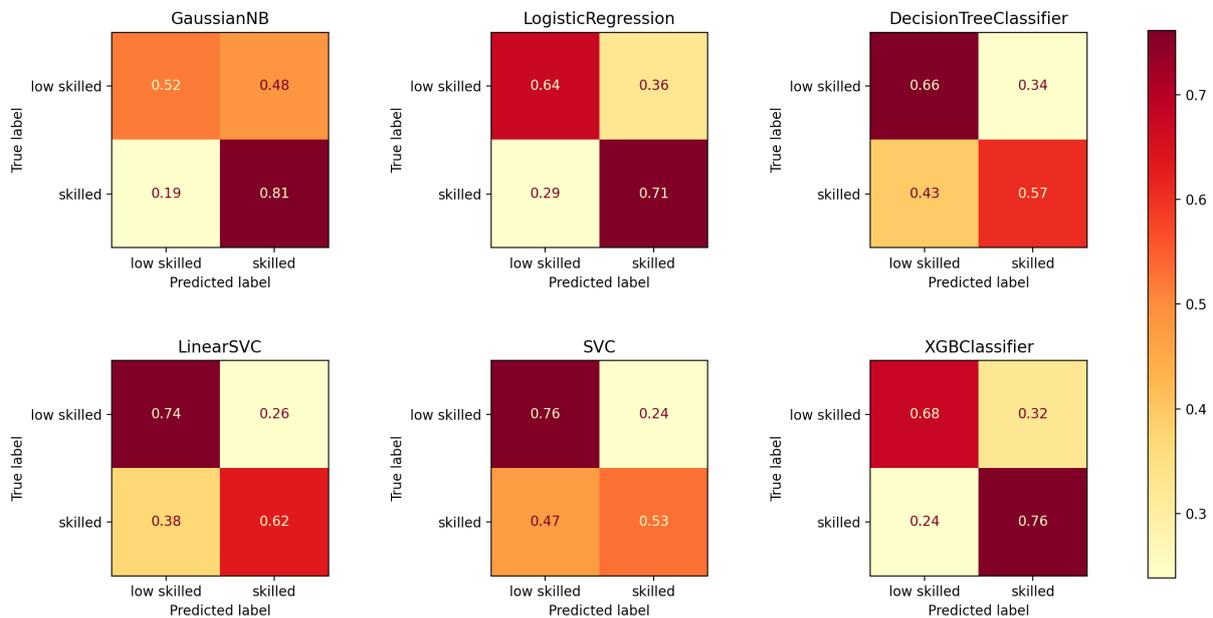


Figure 3: Confusion matrices for runaways being low skilled or skilled with all classifiers
 Source: advertisements in São Paulo newspapers.

In fact, the metrics of Table 3 seen before are only different ways of summarising the outcomes of the confusion matrix. Accuracy, for example, is the proportion of correct predictions (true positives + true

⁴⁶ Kuhn and Johnson, *Applied Predictive Modeling*, ch. 11; Mueller and Guido, *Introduction to Machine Learning*, ch. 5; Chicco and Jurman, "Advantages of the Matthews Correlation."

⁴⁷ Mueller and Guido, *Introduction to Machine Learning*, ch. 5.

negatives) to all predictions (true positives + true negatives + false positives + false negatives) and equals 71 percent for the XGBoost classifier. There is no general rule defining a required value for the metrics presented, as what is regarded as a good score depends on the type of the data and the task involved. For a difficult historical subject such as the runaways' skills, and given the limitations of advertisements pointed out before, the results obtained by the best models in Table 3 seem to be quite good.

Predicted skilled runaways

This final section presents the predictions of the best model (the XGBoost classifier) applied to the group of 2,382 runaways from whom there are no data regarding their work activities. How do these predicted skills compare with those possessed by the runaways with informed occupations in the advertisements? Apart from a general description of predicted values, we highlight the results of features that appeared more relevant in the previous analysis of runaways with occupational data – in particular the outcomes concerning the enslaved persons' characteristics of ethnicity, colour and literacy.

The number of skilled runaways predicted by the best model amounted to 52.3 percent of the total 2,382 runaways with unknown occupations. Such a percentage is even higher than that for the group of 994 enslaved workers (46.6%) with skills informed in the announcements. This result suggests that the bias towards skilled runaways in the advertisements identified earlier was really a characteristic of the source, perhaps even magnified by the best classifier. As shown in Table 4, the regional distribution of skilled runaways of the group with known occupations (n = 994) exhibited 92.7% of all skilled runaways clustered in five regions: Central (38.4%), Capital (25.4%), Paraíba Valley (11.0%), Paulista (9.5%) and Mogiana (8.4%). With reference to the predicted skilled fugitives in the group with unknown occupations (n = 2,382), the same five regions accounted for a similar proportion of skilled runaways (96.5%), with the Central zone (50.6%) also in the top, but now with Paulista in second place (17.5%), ahead of the Capital region (16.9%) (Table 4).

Table 4: Regional distribution, observed and predicted skilled runaways, São Paulo, 1854-1887

Regions (n=994)	%	Regions predicted (n=2,382)	%	Municipalities (n=994)	%	Municipalities predicted (n=2,382)	%
Central	38.4	Central	50.6	Capital	25.2	Campinas	27.1
Capital	25.4	Paulista	17.5	Campinas	15.1	Capital	16.6
Paraíba Valley	11.0	Capital	16.9	Itatiba	4.7	Itatiba	6.5
Paulista	9.5	Paraíba Valley	8.0	Piracicaba	4.3	Limeira	6.3
Mogiana	8.4	Mogiana	3.5	Itu	3.9	Rio Claro	5.4
Total	92.7	Total	96.5	Total	53.2	Total	61.9

Source: advertisements in São Paulo newspapers.

The municipality predicted to have the largest share of skilled fugitives was Campinas, one of the major centres of coffee production at the time, with 27.1 percent of all skilled runaways – substantially higher than what was recorded for this locality (15.1%) in the advertised group with occupations. The Capital of province was predicted as holding 16.6 percent of skilled fugitives, followed by Itatiba (6.5%), Limeira (6.3%) and Rio Claro (5.4%) – all municipalities located in the main coffee areas of São Paulo. This distribution is different from the group of 994 runaways, which after the Capital (25.2%) and Campinas (15.1%) recorded Itatiba (4.7%), Piracicaba (4.3%) and Itu (3.9%) as the localities with the greatest number of skilled fugitives (Table 4).

Most runaways (72.2%) with unknown jobs escaped from rural areas, in a proportion exactly the same

as that observed in the group with informed occupations (72.3%). According to the predictions made by the best model, the majority (76.8%) of skilled fugitives lived in rural areas – and with a percentage even above that found in the group of 994 runaways (68.7%). When seen from their local of residence, predicted runaways in rural areas continued to be relatively less skilled (54.8%) than those living in urban settings (57.0%), although in a lower degree than the group with advertised occupations – which had 42.5 percent of skilled runaways living in farms and 50.9 percent in urban areas. These results seem to support the view that artisanal occupations were rather common in plantations and farms, as noted before.

Runaways were overwhelmingly males (85.9%) in the large dataset lacking skills information, although with a proportion relatively lower than in the group with informed occupations (96.3%). All females were predicted to work in non-artisanal jobs, following the trend already observed in the group with advertised occupations, which recorded only one (0.2%) female runaway working in a artisanal occupation. These figures indicate that artisanal trades were almost an exclusive male occupation in São Paulo as in other such plantation economies.

Skilled runaways were also older than low-skilled slaves, with a mean age of 29 years compared to 26 years, respectively, in the classified group of occupations. These averages are substantially smaller than those observed in the group of 994 runaways, which exhibited a mean age of 31 and 29 years for skilled and low-skilled fugitives, respectively. The median age (26) was also higher among artisans (non-artisans, 24 years) in the predicted sample, such as with the group of known occupations (30 and 27 years, respectively). The non-parametric tests performed suggest that both samples of skilled and low-skilled runaways in the predicted group are statistically different, in the same way as observed in the dataset with known occupations. The interval between 35 and 54 years was again the age class in which the share of artisans (21.8%) was higher than that of non-artisans (14.3%) – compared to 27.8 and 22.0 percent, respectively, for the runaways with advertised occupations. Thus the notion that skilled slaves were of a higher average age, as their jobs required more training, knowledge and experience, also finds support in the group of predicted skilled runaways.

We now turn to three other particularly noteworthy features of the previous analysis of skills among runaways with informed occupations. To facilitate comparisons, Table 5 summarises the main results for both groups of observed and predicted runaways in terms of ethnicity, skin colour and literacy. As noted earlier, with regards to the ethnic origins of runaways with occupations, Africans (“de nação”) had only a slightly lower percentage (44.7%) of artisans than Brazilian-born slaves (47.5%). Besides this, there was no statistically significant association between the skills of African and creole runaways. The idea therefore that Africans were underrepresented in skilled trades because of language or other social barriers was not confirmed for São Paulo runaways.

As regards the group with unknown occupations, 88.8 percent of runaways were Brazilian-born and the best model found out that they made 85.8 percent of all skilled workers. These figures are very close to what was observed in the group of runaways with occupations (88.9% of runaways were Brazilian-born, representing 89.7% of all skilled fugitives). Still, according to the predicted shares of skilled slaves among Africans (70.8%) and creoles (50.4%), runaways constituted an even larger proportion of skilled Africans than was reported for runaways with informed jobs (Table 5). The chi-square test of independence for the predicted group revealed a statistically significant association between samples, suggesting a possible preference for Africans in skilled trades.⁴⁸

The dataset of all runaways without recorded occupations showed a balance between mixed-race (50.3%) and blacks (49.7%), mirroring the shares found in the group with advertised occupations (50.7% and 49.3%, respectively). As indicated in Table 5, more mixed-race (49.5%) than blacks (44.4%) were artisan workers among those with advertised occupations, a difference however that disappears between

⁴⁸ The chi-square statistic was 27.855, rejecting the null hypothesis of no association between the samples.

Table 5: Best model results and tests, observed and predicted skilled runaways, São Paulo, 1854-1887

Features	Observed skilled runaways	Predicted skilled runaways
<i>Ethnicity</i>	(<i>n</i> = 771)	(<i>n</i> = 2,382)
African	44.7	70.8
Creole	47.5	50.4
Chi-square test	no association	association
<i>Skin colour</i>	(<i>n</i> = 901)	(<i>n</i> = 2,382)
Mixed	49.5	52.2
Black	44.4	52.3
Chi-square test	no association	no association
<i>Literacy</i>	(<i>n</i> = 994)	(<i>n</i> = 2,382)
Literate	64.9	67.6
Illiterate	45.6	52.1
Chi-square test	association	no association

Note: see text for definitions.

mixed-race (52.2%) and blacks (52.3%) in the predicted group of skilled runaways. In both cases the allocation of skilled jobs seems not have been related to skin colour, as suggested by the statistical tests of independence.⁴⁹ As pointed out before, the skill gap between *pardos* and *pretos* among São Paulo runaways – whether it ever existed – seems to be relatively small in comparison with similar slave economies in the Americas.

Literacy is the last feature that deserves attention. As seen before, 5.7 percent of the runaways with known occupations were able to read or write, a higher figure than when all runaways are taken into account (2.7%). Artisans, in particular, were clearly more educated (8.0%) than non-artisans (3.8%). These trends were reflected in the distribution of skills between literate and illiterate runaways, as shown in Table 5: 64.9 percent of the former were skilled compared with 45.6 percent of illiterates. Moreover, the chi-square test showed a statistically significant association between fugitives' literacy and skills, indicating that an ability to read and write could be an advantage for enslaved individuals to learn and work in more qualified jobs.

The group of runaways with unknown occupations exhibited a much lower share of literacy: only 34 (1.4%) out of 2,382 enslaved persons were able to read or write. Still, according to the predictions made, the literate runaways in this group continued to show a greater likelihood of being skilled than those uneducated ones — 67.6% against 52.1%, as in Table 5. Nonetheless, the relationship between runaways' literacy and skills does not seem to be statistically significant at this time.⁵⁰ The value of education – even under the dire conditions of slavery – seemed perceptible with the initial results, but looks ambiguous when the group of predicted skills is considered.

Conclusions

When the US Vice-Consul, Christopher C. Andrews, visited the Ibicaba plantation in 1884, he portrayed the slave mechanics as part of a system that employed modern buildings and machinery for different stages of coffee production. Andrews showed no surprise that enslaved workers were able to perform

⁴⁹ The chi-square statistic for the predicted group equalled 0.001, failing to reject the null hypothesis of independence between samples.

⁵⁰ The chi-square statistic was 2.659, therefore not rejecting the null hypothesis of no association between the samples.

skilled jobs as required by the complex activities of coffee processing. Just as in other plantation economies in the Americas, enslaved Africans and their descendants in São Paulo worked in all sorts of tasks, from agricultural and domestic duties to construction and transport, and the artisanal trades were not different. The lack of detailed information about these skilled jobs and workers, however, has helped to downplay their presence in countries like Brazil, especially in the rural landscape of farms and plantations.

The advertisements of runaways are one of the few but richest records left which help to illuminate the actual work put in by enslaved persons in their daily lives. The announcements published in São Paulo newspapers were able, for example, to capture the varied tasks that bonded workers performed in their life cycle, cropping seasons or in their daily work schedule. Even those workers with artisanal skills were often depicted as doing or having done other and sometimes rather different jobs, including working in the field. These multiple occupations revealed in the advertisements show how the work experience in a rural and slave-based society was neither static nor an easy subject to investigate. But what constitutes an advantage of the source is at same time a caveat about their limits. Previous sections of this article have described in detail the conventions and adjustments which were necessary to make the raw data suitable for quantitative analysis.

The initial account of the group of runaways whose jobs were informed in newspapers revealed other problems with the source. The very high proportion of artisans (46.5%) indicates that advertisements were biased toward skilled slaves and were likely not representative of the whole enslaved population. In part, such bias is related to another characteristic of the source – the wide dominance of male slaves among fugitives, a pattern also observed in other slave societies. Still, with the proper caveats, some important conclusions could be drawn from the data.

Skilled slaves were present in great numbers in the rural setting of farms and plantations rather than clustered mainly in urban areas, as is sometimes portrayed. Perhaps surprisingly, Africans were nearly as skilled as Brazilian-born runaways, a result that conflicts with the view that the former faced more hurdles to integrate than creole slaves. In the same fashion, the skill gap between mixed-race and blacks in São Paulo was relatively low in comparison with other plantation economies. Another significant finding was that, despite the very low literacy of the slave population in general, education could have made a difference for the few runaways who managed to learn to read and write, as literate slaves tended to work in more qualified and artisanal occupations.

The detailed individual data of 994 runaways provided by the advertisements were used to predict, using machine-learning classifiers, the skills of 2,382 runaways with unknown occupations. Overall the results of the best model converged with the findings just described. The predicted skilled runaways were mostly male. They tended to be older than their low-skilled counterparts, being widely present in farms and plantations, perhaps even in a higher proportion than indicated by the group with advertised occupations. Relatively, Africans not only may have been almost as skilled as Brazilian-born runaways, as noted before, but even more skilled and ready for artisanal jobs, according to predicted skills. Again, the difference of artisanal skills between *pardos* and *pretos* was small and probably not significant in both observed and predicted groups. Literate runaways were also predicted as having a higher share of skilled jobs, although in this case the difference was not statistically significant when compared to the group with advertised occupations.

These results drawn from advertisements help to understand the little known and often overlooked segment of skilled slaves and their artisanal occupations. The evidence from announcements gives support for the view that the enslaved labour force in São Paulo was at the same level as free workers as regards the skills and ability to work, despite the appalling conditions under which enslaved persons toiled away in the province. Other elements, such as literacy, health and lack of access to land, seem to have been the greatest burden for former slaves and descendants after Abolition in Brazil.

Appendix

Appendix - Number of observations in the runaway datasets

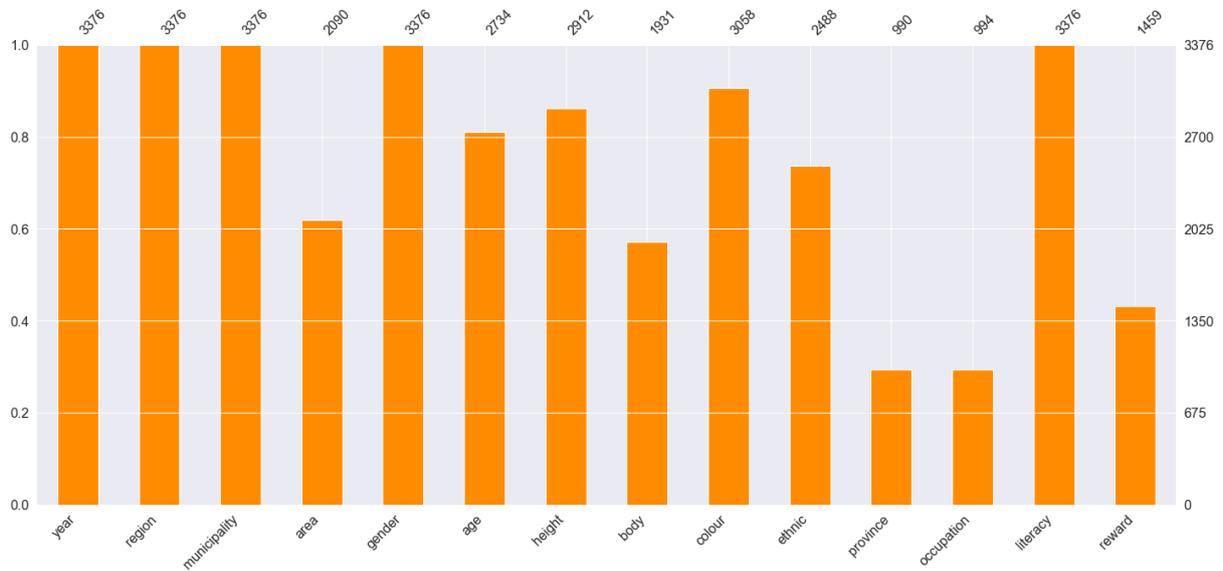


Figure A1: Number of observations for all runaways, São Paulo, 1854-1887

Source: advertisements in São Paulo newspapers.

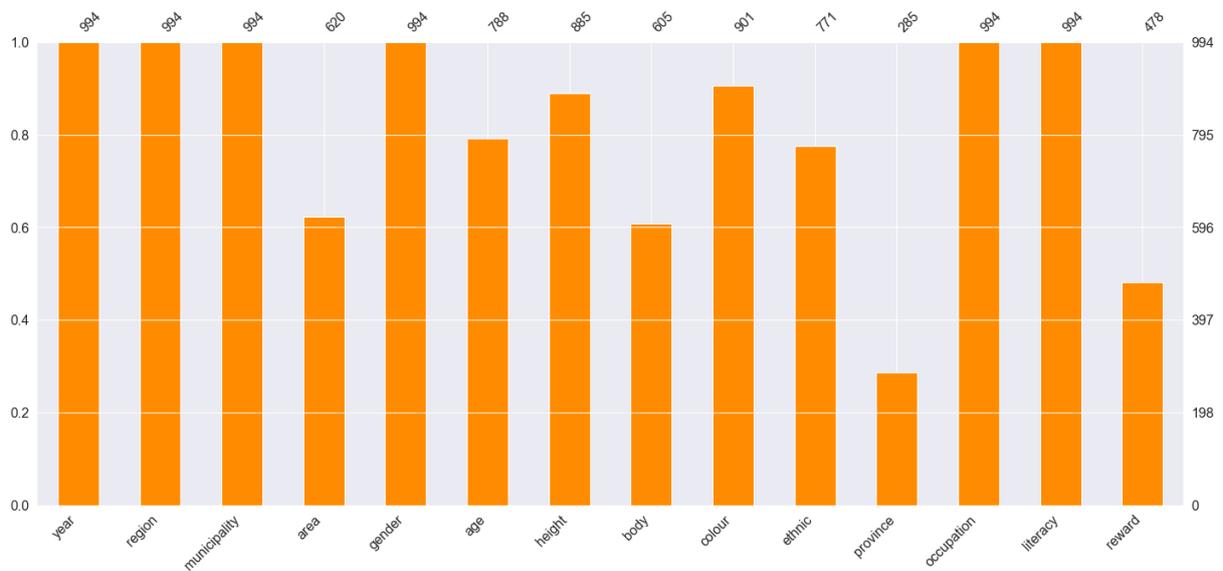


Figure A2: Number of observations for runaways with advertised occupations, São Paulo, 1854-1887

Source: advertisements in São Paulo newspapers.

Appendix - Skilled and low-skilled occupations

Table A1: Main skilled and low-skilled occupations, runaway slaves, São Paulo, 1857-1887

Occupation	Skill classification
Baker	skilled
Barber	low skilled
Blacksmith	skilled
Butcher	low skilled
Butler	low skilled
Cabinet maker	skilled
Carpenter	skilled
Carter	low skilled
Cook	low skilled
Cooper	skilled
Drover	low skilled
Farmer	low skilled
Hatmaker	skilled
Horseman	low skilled
Hunter	low skilled
Laundress	low skilled
Machinist	skilled
Master or official	skilled
Mechanic	skilled
Muleteer	low skilled
Overseer	low skilled
Page	low skilled
Painter	skilled
Ploughman	low skilled
Potter	skilled
Railway worker	skilled
Saddler	skilled
Seamstress	low skilled
Seller (bread, vegetables, others)	low skilled
Shoemaker	skilled
Shop assistant	skilled
Stone mason	skilled
Tailor	skilled
Tanner	skilled
Tiler	skilled
Wagon driver	low skilled
Woodworker, sawyer	low skilled

Source: advertisements in São Paulo newspapers.

References

- Amantino, Marcia. "Os Escravos Fugitivos em Minas Gerais e os Anúncios do Jornal 'O Universal,' 1825 a 1832." *Locus: Revista de História* 12, no. 2 (2006).
- Andrade, Maria José de S. *A Mão de Obra Escrava em Salvador, 1811-1860*. São Paulo: Corrupio, 1988.
- Andrews, Christopher C. *Brazil, Its Conditions and Prospects*. New York: D. Appleton and Company, 1887.
- Barickman, Bert J. "Persistence and Decline: Slave Labour and Sugar Production in the Bahian Reconcavo, 1850-1888." *Journal of Latin American Studies* 28, no. 3 (1996): 581-633.
- Barros, Surya A. P. "Negrinhos que por Ahi Andão: A Escolarização da População Negra em São Paulo (1870-1920)." Master's thesis, University of São Paulo, 2005.
- Baten, Joerg, Ines Pelger, and Linda Twrdek. "The Anthropometric History of Argentina, Brazil and Peru During the 19th and Early 20th Century." *Economics & Human Biology* 7, no. 3 (2009): 319-33.
- Bergad, Laird W. *Slavery and the Demographic and Economic History of Minas Gerais, Brazil, 1720-1888*. Cambridge: Cambridge University Press, 1999.
- Berlin, Ira, and Herbert G. Gutman. "Natives and Immigrants, Free Men and Slaves: Urban Workingmen in the Antebellum American South." *The American Historical Review* 88, no. 5 (1983): 1175-1200.
- Bly, Antonio T. "'Pretends He Can Read': Runaways and Literacy in Colonial America, 1730-1776." *Early American Studies* 6, no. 2 (2008): 261-94.
- Brazil. *Recenseamento Geral do Império de 1872*. Rio de Janeiro: Typ. G. Leuzinger, 1876.
- . *Relatorio e Trabalhos Estatisticos apresentados ao Illm. e Excm. Sr. Conselheiro Dr. João Alfredo Corrêa de Oliveira, Ministro e Secretario d'Estado dos Negocios do Imperio, pelo Director Geral Interino Dr. José Maria do Coutto, em 30 de abril de 1875*. Rio de Janeiro: Typographia de Pinto, Brandão & Comp., 1875.
- Caldas Aulete, J. F. *Diccionario Contemporaneo da Lingua Portuguesa*. 2a ed. Lisboa: Parceria Antonio Maria Pereira, 1925.
- Castro, Hebe M. M. de. *Ao Sul da História*. São Paulo: Editora Brasiliense, 1987.
- Chen, Tianqi, and Carlos Guestrin. "XGBoost: A Scalable Tree Boosting System." In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785-94. San Francisco, CA, 2016.
- Chicco, Davide, and Giuseppe Jurman. "The Advantages of the Matthews Correlation Coefficient (MCC) over F1 Score and Accuracy in Binary Classification Evaluation." *BMC Genomics* 21, no. 1 (2020): 6.
- Colistete, Renato P. "Regiões e Especialização na Agricultura Cafeeira: São Paulo no Início do Século XX." *Revista Brasileira de Economia* 69, no. 3 (2015): 331-54.
- Conrad, Robert E. *Os Últimos Anos da Escravatura no Brasil, 1850-1888*. Rio de Janeiro: Civilização Brasileira, 1975.
- Cornelius, Janet D. *When I Can Read My Title Clear: Literacy, Slavery and Religion in the Antebellum South*. Columbia: University of South Carolina Press, 1991.
- Costa, Ana Caroline de R. "Fugas de Escravos na Comarca do Rio das Mortes, Primeira Metade do Século XIX." Master's thesis, Federal University of São João del Rey, 2013.

- Couty, Louis. *Étude de Biologie Industrielle sur le Café. Rapport adressé a M. le Directeur de l'École Polytechnique*. Rio de Janeiro: Imprimerie du "Messager du Brésil", 1883.
- Cowling, Camillia. *Conceiving Freedom: Women of Color, Gender, and the Abolition of Slavery in Havana and Rio de Janeiro*. Chapel Hill: University of North Carolina Press, 2013.
- Craton, Michael. *Searching for the Invisible Man: Slaves and Plantation Life in Jamaica*. Cambridge, Mass: Harvard University Press, 1978.
- Crüwell, G. A. *Brazil as a Coffee-Growing Country: Its Capabilities, the Mode of Cultivation, and Prospects of Extension*. Colombo: A. M. & J. Ferguson, 1878.
- Dean, Warren. *Rio Claro: A Brazilian Plantation System, 1820-1920*. Stanford: Stanford University Press, 1976.
- Degler, Carl. *Neither Black Nor White: Slavery and Race Relations in Brazil and the United States*. Madison: University of Wisconsin Press, 1986.
- Dias, Maria Odila da Silva. "Nas Fímbricas da Escravidão Urbana." *Estudos Econômicos* 15 (1985): 89–109.
- Dittmar, Jeremiah, and Suresh Naidu. "Contested Property: Fugitive Slaves in the Antebellum U.S. South." Manuscript, 2012.
- Dunn, Richard S. *A Tale of Two Plantations: Slave Life and Labor in Jamaica and Virginia*. Cambridge, Mass.: Harvard University Press, 2014.
- . *Sugar and Slaves: The Rise of the Planter Class in the English West Indies, 1624-1713*. Second. Chapel Hill: University of North Carolina Press, 2000.
- Efron, Bradley, and Trevor Hastie. *Computer Age Statistical Inference: Algorithms, Evidence, and Data Science*. Cambridge: Cambridge University Press, 2016.
- Fallope, Josette. "Les Occupations d'Esclaves à la Guadeloupe dans la Première Moitié du XIXe Siècle." *Outre-Mers. Revue d'histoire* 74, no. 275 (1987): 189–205.
- Feinstein, Charles H., and Mark Thomas. *Making History Count: A Primer in Quantitative Methods for Historians*. Cambridge: Cambridge University Press, 2002.
- Ferlini, Vera Lúcia Amaral. *Terra, Trabalho e Poder: O Mundo dos Engenhos no Nordeste Colonial*. Florianópolis: EDUSC, 2003.
- Florentino, Manolo, and José Roberto Góes. "Parentesco, Estabilidade Familiar e Ocupação entre os Escravos: O Caso do Meio Rural Fluminense entre 1790 e 1830." *Revista Brasileira de Estudos de População* 13, no. 2 (1996): 183–98.
- Fogel, Robert William. *Without Consent or Contract: The Rise and Fall of American Slavery: The Rise and Fall of American Slavery*. New York: W. W. Norton & Company, 1989.
- Fogel, Robert William, and Stanley L Engerman. *Time on the Cross. The Economics of American Negro Slavery*. Boston: Little, Brown and Company, 1974.
- Foner, Eric. *Gateway to Freedom. The Hidden History of the Underground Railroad*. Oxford: Oxford University Press, 2015.
- Fraga, Walter. *Crossroads of Freedom: Slaves and Freed People in Bahia, Brazil, 1870-1910*. Durham: Duke University Press, 2016.
- Fragoso, João Luis R. "Sistemas Agrários em Paraíba do Sul (1850-1920): Um Estudo de Relações Não-Capitalistas de Produção." Master's thesis, Federal University of Rio de Janeiro, 1983.

- Franken, Daniel. "Anthropometric History of Brazil, 1850–1950: Insights from Military and Passport Records." *Revista de Historia Económica - Journal of Iberian and Latin American Economic History* 37, no. 2 (2019): 377–408.
- Franklin, John Hope, and Loren Schwenger. *Runaway Slaves: Rebels on the Plantation*. Oxford: Oxford University Press, 2000.
- Freyre, Gilberto. *O Escravo nos Anúncios de Jornais Brasileiros do Século XIX*. 2nd ed. São Paulo: Companhia Editora Nacional, 1979.
- Gebara, Ademir. "Escravidão: Fugas e Controle Social." *Estudos Econômicos* 18 (1988): 103–46.
- Godoy, Marcelo M. "Fazendas Diversificadas, Escravos Polivalentes. Caracterização Sócio-demográfica e Ocupacional dos Trabalhadores Cativos em Unidades Produtivas com Atividades Agroaçuareiras de Minas Gerais no século XIX." In *Anais do XIV Encontro Nacional de Estudos Populacionais*. Caxambú-MG, 2004.
- Gomes, Flávio. "A Demografia Atlântica dos Africanos no Rio de Janeiro, Séculos XVII, XVIII e XIX: Algumas Configurações a partir dos Registros Eclesiásticos." *História, Ciências, Saúde-Manguinhos* 19 (2012): 81–106.
- . "Jogando a Rede, Revendo as Malhas: Fugas e Fugitivos No Brasil Escravista." *Tempo* 1 (1996): 67–93.
- Graham, Richard. "Slave Families of a Rural Estate in Colonial Brazil." *Journal of Social History* 9, no. 3 (1976): 382–402.
- Graham, Sandra L. "Writing from the Margins: Brazilian Slaves and Written Culture." *Comparative Studies in Society and History* 49, no. 3 (2007): 611–36.
- Hall, Douglas. "Slaves and Slavery in the British West Indies." *Social and Economic Studies* 11, no. 4 (1962): 305–18.
- Hardesty, Jared Ross. "'The Negro at the Gate': Enslaved Labor in Eighteenth-Century Boston." *The New England Quarterly* 87, no. 1 (2014): 72–98.
- Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. New York: Springer, 2009.
- Heuman, Gad. "Runaway Slaves in Nineteenth-Century Barbados." *Slavery & Abolition* 6, no. 3 (1985): 95–111.
- Higman, B. W. *Slave Populations of the British Caribbean, 1807-1834*. University of the West Indies Press, 1995.
- Holanda, Sérgio Buarque de. "Prefácio." In *Memórias de Um Colono No Brasil (1850)*. São Paulo: Martins Fontes, 1941.
- Hunter, John D. "Matplotlib: A 2d Graphics Environment." *Computing in Science & Engineering* 9, no. 3 (2007): 90–95.
- Irwin, James R. "Farmers and Laborers: A Note on Black Occupations in the Postbellum South." *Agricultural History* 64, no. 1 (1990): 53–60.
- Johnson, Michael P. "Runaway Slaves and the Slave Communities in South Carolina, 1799 to 1830." *The William and Mary Quarterly* 38, no. 3 (1981): 418–41.
- . "Work, Culture, and the Slave Community: Slave Occupations in the Cotton Belt in 1860." *Labor History* 27, no. 3 (1986): 325–55.

- Jordahl, Kelsey, Joris Van den Bossche, Martin Fleischmann, and others. "GeoPandas: Python Tools for Geographic Data." Zenodo, 2020.
- Karasch, Mary C. *Slave Life in Rio de Janeiro, 1808-1850*. Princeton, N.J.: Princeton University Press, 1987.
- Kuhn, Max, and Kjell Johnson. *Applied Predictive Modeling*. New York: Springer, 2013.
- Kulikoff, Allan. *Tobacco and Slaves: The Development of Southern Cultures in the Chesapeake, 1680-1800*. Chapel Hill: University of North Carolina Press, 1986.
- Laërne, C. F. Van Delden. *Brazil and Java: Report on Coffee-Culture in America, Asia and Africa, to H.E. The Minister of the Colonies*. London: W.H. Allen & Company, 1885.
- Lago, Luiz Aranha C. do. *Da Escravidão ao Trabalho Livre*. São Paulo: Companhia das Letras, 2014.
- Lamberg, Mauricio. *O Brazil*. Rio de Janeiro: Typographia Nunes, 1896.
- Lima, Carlos A. M. *Artífices do Rio de Janeiro*. Rio de Janeiro: Apicuri, 2008.
- . "Escravos Artesãos: Preço e Família. Rio de Janeiro, 1789-1839." *Estudos Econômicos* 30, no. 3 (2000): 447-84.
- Linsin, Christopher E. "Skilled Slave Labor in Florida: 1850-1860." *The Florida Historical Quarterly* 75, no. 2 (1996): 183-96.
- Lobo, Eulalia Maria Lahmeyer, Otávio Canavarros, Zakia Feres Elias, Simone Novais, and Lucena Barbosa Madureira. "Estudo das Categorias Socioprofissionais, dos Salários e do Custo da Alimentação no Rio de Janeiro de 1820 a 1930." *Revista Brasileira de Economia* 27, no. 4 (1973): 129-76.
- Lowrie, Samuel H. "The Negro Element in the Population of Sao Paulo, a Southernly State of Brazil." *Phylon* 3, no. 4 (1942): 398-416.
- Machado, Geosiane M. "Com Vistas à Liberdade: Fugas Escravas e Estratégias de Inserção Social do Fugido nos Últimos Décênios do Século XIX, em Minas Gerais." Master's thesis, Federal University of Minas Gerais, 2010.
- Mair, Lucille M. *A Historical Study of Women in Jamaica, 1655-1844*. Edited by Hilary McD Beckles and Verene A. Shepherd. Kingston: University Press of the West Indies, 2007.
- Marks, Bayly E. "Skilled Blacks in Antebellum St. Mary's County, Maryland." *The Journal of Southern History* 53, no. 4 (1987): 537-64.
- Mattos, Hebe. *Das Cores do Silêncio*. 3ª edição. Rio de Janeiro: Editora da Unicamp, 2013.
- Mattos, Regiane A. "De Cassange, Mina, Benguela a Gêntio da Guiné: Grupos Étnicos e Formação de Identidades Africanas na Cidade de São Paulo (1800-1850)." Master's thesis, University of São Paulo, 2006.
- McGinnis, William D., Chapman Siu, Andre S, and Hanyu Huang. "Category Encoders: A Scikit-Learn-Contrib Package of Transformers for Encoding Categorical Data." *Journal of Open Source Software* 3, no. 21 (2018): 501.
- McKinney, Wes. "Data Structures for Statistical Computing in Python." In *Proceedings of the 9th Python in Science Conference*, 56-61. Austin, Texas, 2010.
- Menard, Russell R. "The Maryland Slave Population, 1658 to 1730: A Demographic Profile of Blacks in Four Counties." *The William and Mary Quarterly* 32, no. 1 (1975): 29-54.
- Milliet, Sérgio. *Roteiro do Café e Outros Ensaios*. 4th ed. São Paulo: Hucitec, 1982.

- Morais, Christianni C. "Ler e Escrever: Habilidades de Escravos e Forros? Comarca do Rio das Mortes, Minas Gerais, 1731-1850." *Revista Brasileira de Educação* 12, no. 36 (2007): 493-504.
- Mueller, Andreas C., and Sarah Guido. *Introduction to Machine Learning with Python: A Guide for Data Scientists*. Sebastopol, CA: O'Reilly, 2016.
- Nishida, Mieko. "Manumission and Ethnicity in Urban Slavery: Salvador, Brazil, 1808-1888." *Hispanic American Historical Review* 73, no. 3 (1993): 361-91.
- Oliveira, Joice F. "Forasteiros no Oeste Paulista: Escravos no Comércio Interno de Cativos e suas Experiências em Campinas, 1850-1888." Master's thesis, State University of Campinas, 2013.
- Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, et al. "Scikit-Learn: Machine Learning in Python." *Journal of Machine Learning Research* 12, no. 85 (2011): 2825-30.
- Penna, Clemente. "Escravidão, Liberdade e os Arranjos de Trabalho na Ilha de Santa Catarina nas Últimas Décadas de Escravidão (1850-1888)." Master's thesis, Federal University of Santa Catarina, 2005.
- Piza, Edith, and Fúlvia Rosemberg. "Cor nos Censos Brasileiros." *Revista USP*, no. 40 (1999): 122-37.
- Pritchett, Jonathan, and Jessica Hayes. "The Occupations of Slaves Sold in New Orleans: Missing Values, Cheap Talk, or Informative Advertising?" *Cliometrica* 10, no. 2 (2016): 181-95.
- Prude, Jonathan. "To Look Upon the 'Lower Sort': Runaway Ads and the Appearance of Unfree Laborers in America, 1750-1800." *The Journal of American History* 78, no. 1 (1991): 124-59.
- Read, Ian. *The Hierarchies of Slavery in Santos, Brazil, 1822-1888*. Stanford: Stanford University Press, 2012.
- Read, Ian, and Kari Zimmerman. "Freedom for Too Few: Slave Runaways in the Brazilian Empire." *Journal of Social History* 48, no. 2 (2014): 404-26.
- Reis, João José. "Identidade e Diversidade Étnicas nas Irmandades Negras no Tempo da Escravidão." *Tempo* 2, no. 3 (1996): 7-33.
- Rey, Sergio J., and Luc Anselin. "PySAL: A Python Library of Spatial Analytical Methods." *The Review of Regional Studies* 37, no. 1 (2007): 5-27.
- Roberts, Justin. "Working Between the Lines: Labor and Agriculture on Two Barbadian Sugar Plantations, 1796-97." *The William and Mary Quarterly* 63, no. 3 (2006): 551-86.
- Rodarte, Mario Marcos Sampaio, and José Maria dos Santos Júnior. "A Estrutura Ocupacional Revisitada: Uma Proposta de Correção dos Dados do Recenseamento Geral do Império de 1872." In *Anais do XVI Encontro Nacional de Estudos Populacionais*, 0:1-21. Caxambú-MG, 2008.
- Santos, Martha S. "Mothering Slaves, Labor, and the Persistence of Slavery in Northeast Brazil: A Non-Plantation View from the Hinterlands of Ceará, 1813-1884." *Women's History Review* 27, no. 6 (2018): 954-71.
- Schwartz, Stuart B. "A Commonwealth Within Itself. The Early Brazilian Sugar Industry, 1550-1670." In *Tropical Babels. Sugar and the Making of the Atlantic World, 1450-1680*, edited by Stuart B. Schwartz, 158-200. Amsterdam: Amsterdam University Press, 2004.
- . *Sugar Plantations in the Formation of Brazilian Society: Bahia, 1550-1835*: 52. Cambridge: Cambridge University Press, 1985.

- Schweninger, Loren. "Counting the Costs: Southern Planters and the Problem of Runaway Slaves, 1790-1860." *Business and Economic History* 28, no. 2 (1999): 267-76.
- Sheridan, Richard B. "Changing Sugar Technology and the Labour Nexus in the British Caribbean, 1750-1900, with Special Reference to Barbados and Jamaica." *Nieuwe West-Indische Gids / New West Indian Guide* 63, no. 1/2 (1989): 59-93.
- Sidbury, James. "Slave Resistance." In *The Routledge History of Slavery*, edited by Gad Heuman and Trevor Burnard. Oxford: Routledge, 2011.
- Siegel, Sidney. *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill, 1956.
- Silva Pinto, Luiz Maria da. *Dicionário da Língua Brasileira*. Ouro Preto: Typographia de Silva, 1832.
- Slenes, Robert W. "'Malungu, Ngoma Vem!': África Coberta e Descoberta do Brasil." *Revista USP*, no. 12 (1992): 48-67.
- . "The Brazilian Internal Slave Trade, 1850-1888: Regional Economies, Slave Experience, and the Politics of a Peculiar Market." In *The Chattel Principle: Internal Slave Trades in the Americas*, edited by Walter Johnson, 325-70. New Haven: Yale University Press, 2004.
- . "The Demography and Economics of Brazilian Slavery, 1850-1888." PhD thesis, Stanford University, 1975.
- Soares, Luiz Carlos. *O "Povo de Cam" na Capital do Brasil - A Escravidão Urbana no Rio de Janeiro do Século XIX*. Rio de Janeiro: 7 Letras, 2008.
- Stein, Stanley J. *Vassouras. A Brazilian Coffee County, 1850-1900. The Roles of Planter and Slave in a Plantation Society*. Princeton, N.J.: Princeton University Press, [1957] 1992.
- Van Rossum, Guido, and Fred L. Drake. *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace, 2009.
- Virtanen, Paul, Ralf Gommers, Travis E. Oliphant, and others. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python." *Nature Methods* 17, no. 3 (2020): 261-72.
- Waldstreicher, David. "Reading the Runaways: Self-Fashioning, Print Culture, and Confidence in Slavery in the Eighteenth-Century Mid-Atlantic." *The William and Mary Quarterly* 56, no. 2 (1999): 243-72.
- Wallace, Shaun. "Fugitive Slave Advertisements and the Rebelliousness of Enslaved People in Georgia and Maryland, 1790-1810." PhD thesis, University of Stirling, 2017.
- Ward, J. R. *British West Indian Slavery, 1750-1834: The Process of Amelioration*. Oxford: Oxford University Press, 1988.
- White, Shane. "Black Fugitives in Colonial South Carolina." *Australia New Zealand American Studies Association* 1, no. 1 (1980): 25-40.
- Wissenbach, Maria Cristina C. "Arranjos da Sobrevivência Escrava na Cidade de São Paulo do Século XIX." *Revista de História*, no. 119 (1988): 101-13.
- Witzel de Souza, Bruno G. "From Bonded Laborers to Educated Citizens? Immigration, Labor Markets, and Human Capital in São Paulo, Brazil (1820-2010)." PhD thesis, Georg-August-Universität Göttingen, 2019.