# WGBH - Color of Money

Elisa, Jena, Murt, Richard, Sabrina CS506 - Spring 2021



# **Background and Motivation**

- Massachusetts construction projects
- Summary of Workforce Utilization Report
- Minorities rate between 2019 and 2020









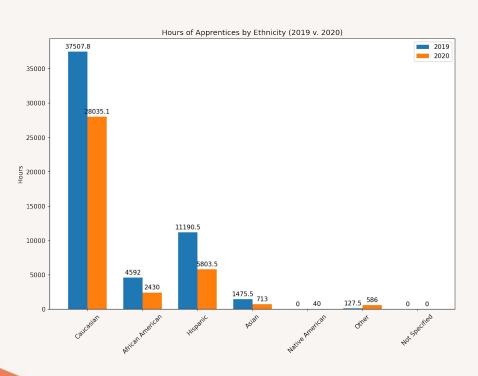
Create an universal parser

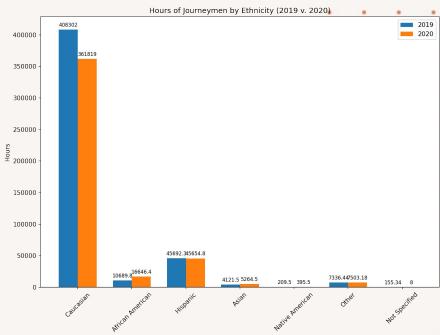


Analyse the minorities rates



## **General Trends**

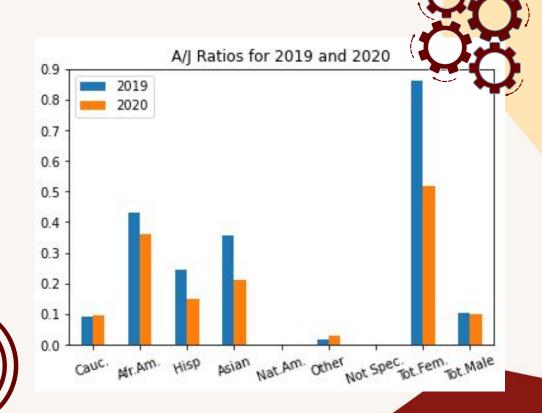


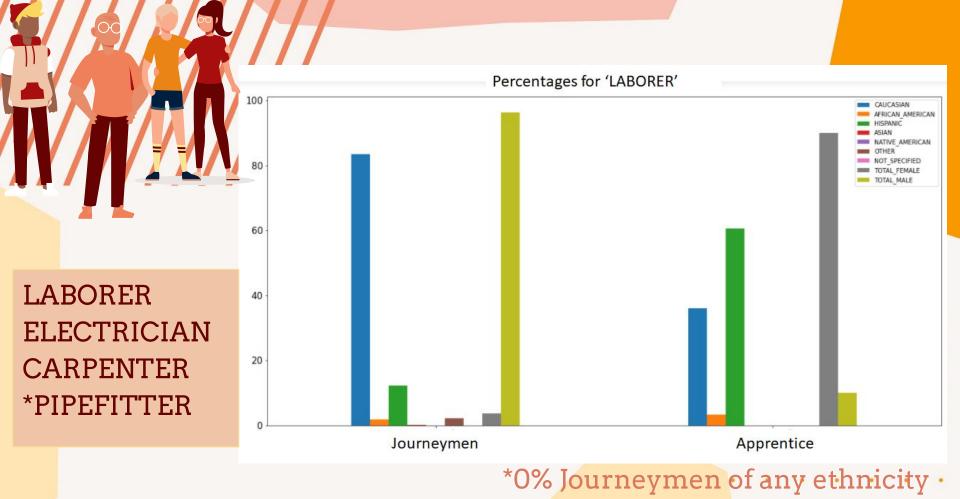


Apprentice to Journeymen Differences

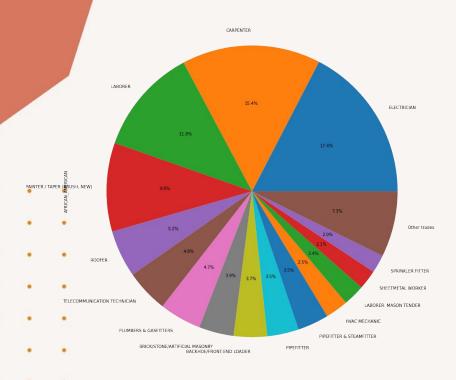
**32%** of all **JOURNEYMEN** hours have at least 95% going to **CAUCASIANS** (2019)

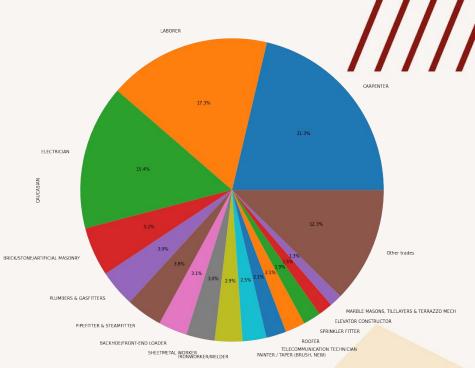
7% of all APPRENTICE hours have at least 95% going to CAUCASIANS (2019)











# Challenges



### **Data Format** Challenges



#### Report Header Separate DF: (3 rows & two categories)

Trade Totals (5 lines):

Contractor Totals (5 lines):



#### Contractor Header (1 lines):

#### 33,000 lines of text and 1,097 pages **Data Format** 2019 Data LIMITATIONS Challenges We need more years for trend analysis Report Header Project Code: CHE1604 DC1 CM (3 rows & two categories) Project Totals (4 lines): Mar2819 pages 25 Apr2819 pages 82 May2819 pages 89 PG 2 of 82 Jun2819 pages 98 Sep2819 pages 183 Dot2819 pages 188 Project End (1 lines): Nov2819 pages 129

#### **CHALLENGES**

- Directly read the PDF file into a Pandas DataFrame, each pg .. df
- Manually transformed into CSV
- PyPDF2, Tabula, Camelot, Pandas, PDFplumber, and batch Tabula-CSV



# Truth in Publishing

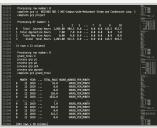
## **News Stories**

If WGBH wants to use the information we identify then they need proof that we parsed all that data according to the PDFs they received.

See excerpts from proof.txt:

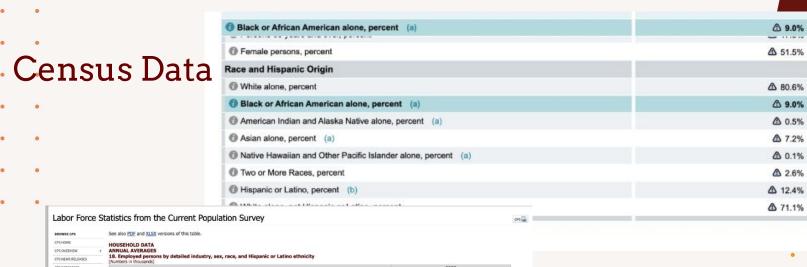
FYI, proof.txt is 115,304 lines long.







## **Limitations**: Did You Know This?



PSHOME	HOUSEHOLD DATA						
CPS OVERVIEW #	ANNUAL AVERAGES  18. Employed persons by detailed industry, sex, race, and Hispanic or Latino ethnicity						
CPS NEWS RELEASES	Numbers in thousands						
OPS DATABASES	2020						
OPS TABLES		Total employed	Percent of total employed				
OPS PUBLICATIONS	Industry		Women	Afr	Black or		Hispani or Latino
OPS FAQS					African American	Asian	
CONTACT CPS	Total, 16 years and over	147,795	46.8	78.0	12.1	6.4	17
SEARCH CPS	Agriculture, forestry, fishing, and hunting	2,349	27.7	92.4	2.8		24
	Crop production	1,186	28.9	92.7	2.2	1.7	30
	Animal production and aquaculture	832	28.1	92.9	3.3	0.4	17
CPS TOPICS	Support activities for agriculture and forestry	145	33.9	90.4	5.7		26
OPS TOPICS A TO Z	Forestry, except logging	53	25.5	97.8			5
25 TOPICS A TO Z	Logging	88	7.3	89.9	3.8	2.0	12
DEMOGRAPHICS >	Fishing, hunting, and trapping	45					
EMPLOYMENT >	Mining, quarrying, and oil and gas extraction	684	14.5	87.2		4.2	1
INEMPLOYMENT #	Oil and gas extraction	97	14.9	80.6		10.5	
ARNINGS F	Coal mining	51	10.1	93.1	3.1	0.0	
	Metal ore mining	36		-			
HOURS OF WORK >	Nonmetallic mineral mining and quarrying	83	10.7	90.4	8.0	2.5	1
TECHNICAL DOCUMENTATION:	Not specified type of mining	0	-			-	
	Support activities for mining	417	15.2	86.5	6.0	3.9	20
RELATED LINKS	Construction	10,786	10.9	88.6	6.0	2.0	30
INFORMATION FOR CPS PARTICIPANTS	Manufacturing	14,550	29.5	79.7	10.3	7.0	1
	Durable goods manufacturing	9,129	25.4	80.3	9.5		15
STATE AND LOCAL LABOR FORCE DATA	Nonmetallic mineral products manufacturing	410	21.3	84.0	8.8	2.4	2
	Pottery, ceramics, and plumbing fixture manufacturing	17		-		-	
NONFARM PAYROLL EMPLOYMENT	Clay building material and refractories manufacturing	38					
	Glass and glass product manufacturing	123	22.9	78.9			2
	Cement, concrete, lime, and gypsum product manufacturing	157	12.9	89.3			2
	AND	70	20.6	DA P	70	10	- 1

## **Next Steps**



375000

## WHAT IS THE CAUSE?

WHY?

The contracting office has influence?
The location influences available workforce?
The companies' influence outcomes?
Friendliness of trades to ethnicities and women?

INTERESTING OBSERVATIONS

Apprentice and Journeymen Utilization.

How many people of color live in MA?

We found interesting patterns in

## WHAT FACILITATES ANALYSIS?

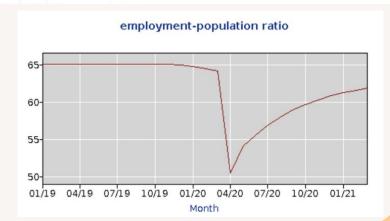
Additional data source with zip code data.

## Resources

#### Earnings by Occupation

	Wages, 2020						
	Hour	Annual					
Data series	Median	Mean	Median	Mean			
Carpenters	\$23.91	\$26.19	\$49,730	\$54,470			
Construction laborers	\$18.38	\$20.92	\$38,220	\$43,520			
Construction managers	\$46.09	\$51.02	\$95,870	\$106,120			
Electricians	\$26.62	\$29.23	\$55,360	\$60,800			
Operating engineers and other construction equipment operators	\$25.04	\$28.11	\$52,090	\$58,470			

(Source: Occupational Employment and Wage Statistics)



- INTRODUCTION AND OVERVIEW OF METHODOLOGY
- Tables
- List of Reports
- Report No. 1 MBE and WBE Payments for Construction Projects FY2020
- Report No. 2 MBE and WBE Payments for Design Projects FY2020
- Report No. 3 MBE, WBE, and Veteran Prime Payments on DCAMM Projects FY2020
- Report No. 4 MBE Payments Distribution in Construction FY2020
- Report No. 5 WBE Payments Distribution in Construction FY2020
- Report No. 6 MBE Payments Distribution in Design FY2020
- Report No. 7 WBE Payments Distribution in Design FY2020
- Report No. 8 MBE and WBE Construction Project Awards FY2020
- Report No. 9 MBE and WBE Design Project Awards FY2020
- Report No. 10 MBE Awards Distribution in Construction Projects FY2020
- Report No. 11 WBE Awards Distribution in Construction Projects FY2020
- Report No. 12 MBE Awards Distribution in Design Projects FY2020
- Report No. 13 WBE Awards Distribution in Design Projects FY2020
- Report No. 14 Construction Hours of Employment by ZIP Code FY2020
- Report No. 15 Design Hours of Employment by ZIP Code FY2020

# Thanks!

WGBH - DCAMM's TEAM

CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon, and infographics & images by Freepik.

Please keep this slide for attribution.