Unemployment Status of the Civilian Population

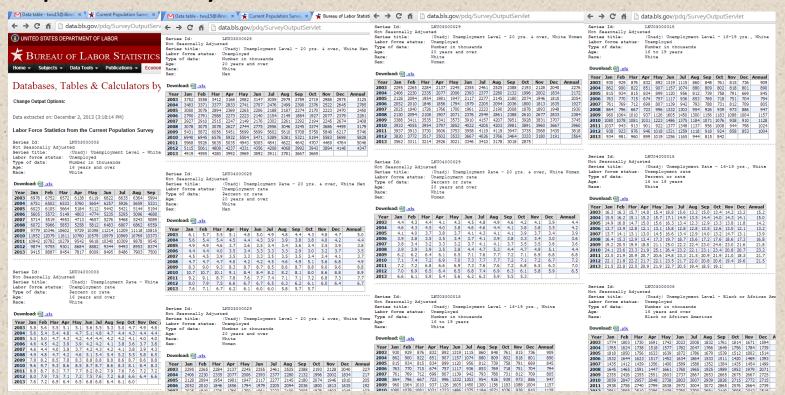
Tai Lin Wu & Brian Winkelmann

Overview:

Discovering the correlation and patterns among race, sex, and age in unemployment data

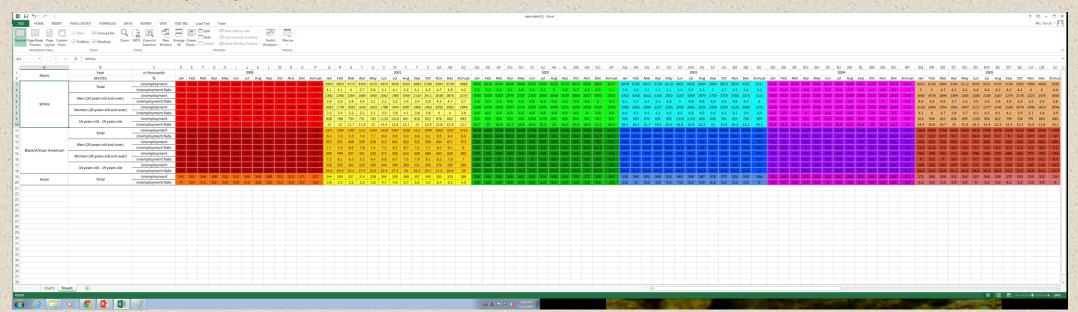
Data Collection

 Use unemployment data that is provided by Bureau of Labor Statistics



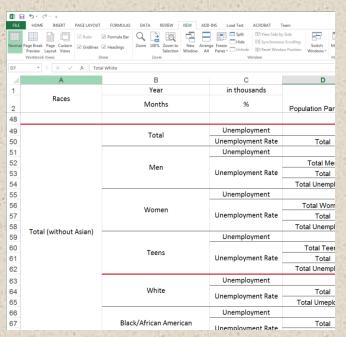
Data Collection

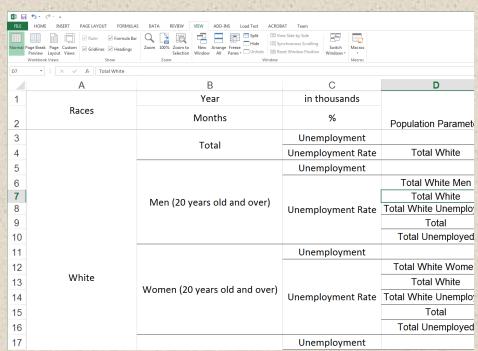
- We organized the data into one excel file to then we split the data up by categories race, sex, and age
- Also split up based on unemployment rate or unemployed number



Data Collection

- Then we bring into Matlab
- Through Matlab, we generated even more data, such as Unemployment rate with total population (unemployed + employed), total unemployed population, Total population in certain race...etc.





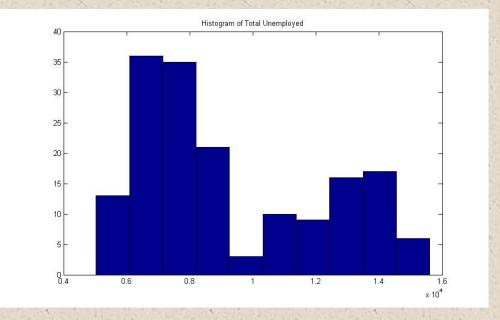
Data collection

After having all the useful data, we then calculated mean,
SD, covariance.

GE	GF	GG	GH	GI	GJ	GK	GL	GM	GN
Overall					Annual				
mean	variance	standard deviation	covarianc e	correlation coefficient	mean	variance	standard deviation	covarianc e	correlation coefficient
4.887952	2.119734	1.455931			4.823077	2.026391	1.423513		
2.123685	0.421469	0.649206			2.092999	0.406926	0.637908		
37.592416	5.983397	2.446098			37.477136	1.606912	1.267641		
1.860701	0.320266	0.565921	0.017389	0.015665	1.833300	0.310225	0.556979	0.017149	0.015884
29.216063	2.761522	1.661783	-0.008056	-0.002346	29.167176	0.578128	0.760347	-0.008011	-0.005923
931.500000	32820.924699	181.165462			928.307692	13396.828402	115.744669		
16.781928	17.044131	4.128454			16.546154	15.108639	3.886983		
0.757546	0.019363	0.139150			0.755188	0.007381	0.085913		
14.151468	10.533415	3.245522			14.283074	7.123627	2.669012		
0.663952	0.014755	0.121471	0.006217	0.026495	0.661527	0.005642	0.075115	0.006200	0.048251
10.982219	5.887337	2.426384	0.000125	0.000025	11.101557	3.953977	1.988461	0.000061	0.000020
1969.819277	308763.593845	555.665001			1937.000000	292662.000000	540.982440		
11.302410	8.133729	2.851969			11.176923	7.829467	2.798119		
878.313253	95395.419945	308.861490			863.538462	92560.248521	304.237158		
11.229518	12.611960	3.551332			11.115385	12.279763	3.504249		
5.031585	2.659193	1.630703			4.972520	2.584639	1.607681		
43.883490	17.858110	4.225886			43.777440	9.919632	3.149545		
0.624323	0.044538	0.211040	0.005813	0.013672	0.613205	0.042785	0.206845	0.005716	0.014105
9.718892	0.625922	0.791152	-0.007932	-0.001203	9.669576	0.221041	0.470150	-0.007953	-0.002250
830.753012	56950.969117	238.644022			814.538462	50850.710059	225.501020		

Data Limitations

- An Education category would have been nice to do more comparisons
- More data points could help the distributions to be more normal
- The Asian categories was limited
 - It was not split up into Men, Women, and Teens
- The race categories were restricted
 - Could use a Hispanic category



Data Analysis

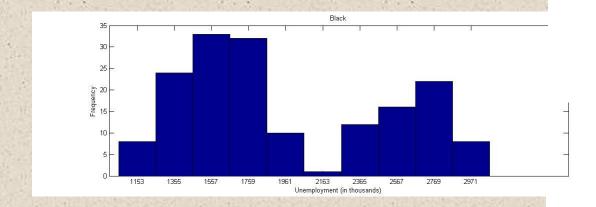
- Get mean and variance for each category
- Plot the distributions and find any interesting patterns
- Discovering the correlation among the categories
 - We got the covariance between two categories
 - Cov(x, y) = E[xy] E[x]E[y]
 - Ex. Cov(White, Men) = E[White Men] E[White]E[Men]
- With covariance we can see if categories are independent

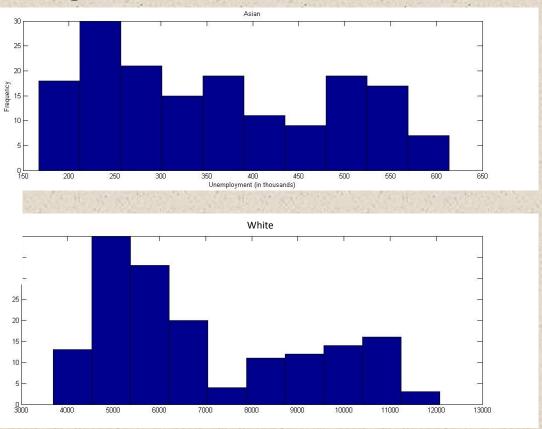
Data Analysis

- For each category:
 - Unemployment rate = Unemployment number/ Total population
 - Total population (Unemployment + Employment) = Total unemployment / Unemployment Rate
- Do hypothesis testing on Unemployment Rate
 - Comparing two distinct populations
 - Using ML decision rule
 - Find p_{false-alarm} and p_{miss}

Results

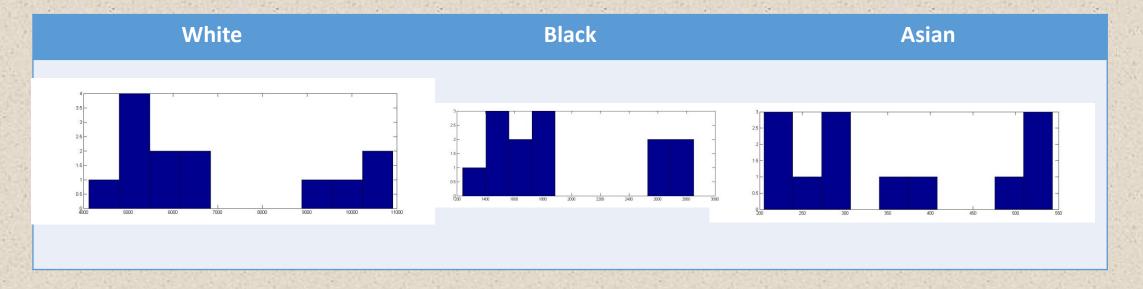
- We took each set and plotted their histogram
- They don't look normal





Annual Data

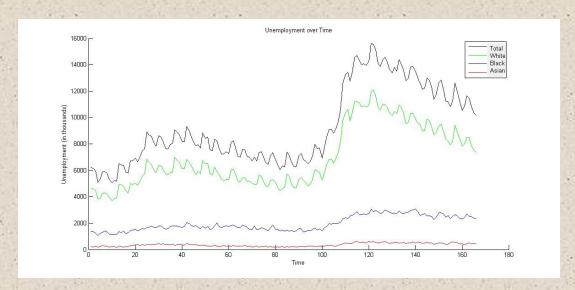
- The mean of the Annual data is less than mean of the Monthly data
- The covariance using Annual data is close the covariance using Monthly data
- The Annual data shows the gap between the two peeks better

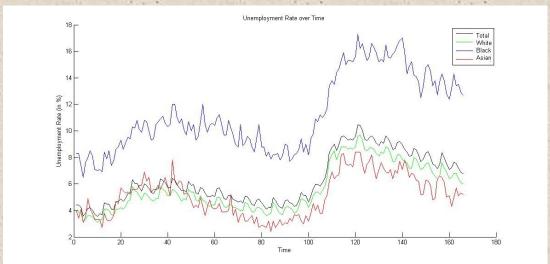


White, Black, and Asian

This plot shows each distribution follow a similar curve over time

Comparing Total, White, Black, and Asian Unemployment Rate

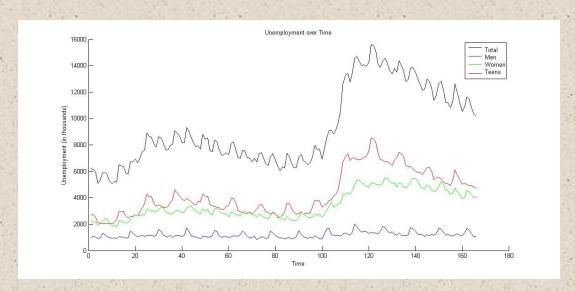


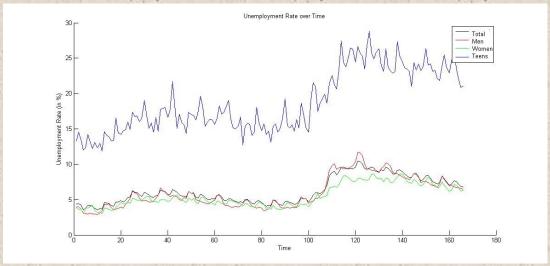


Men, Women, and Teens

Once again we can see that they follow the same curve

Comparing Total, Men, Women, and Teens Unemployment Rate





Summary

- We were able to take the unemployment data and ...
 - We find the covariance for each category is never 0, meaning none of them is completely independent to others.
 - However covariance we found are all really small, meaning they are not very correlated
 - We also found some negative covariance, meaning some of the categories, such as are negatively correlated

Improve Projects

- Written outlines
 - Instructions that are less vague
 - Make clear the difficult parts
- Have the two parts more equal in workload
 - Have a Matlab file already with the data imported