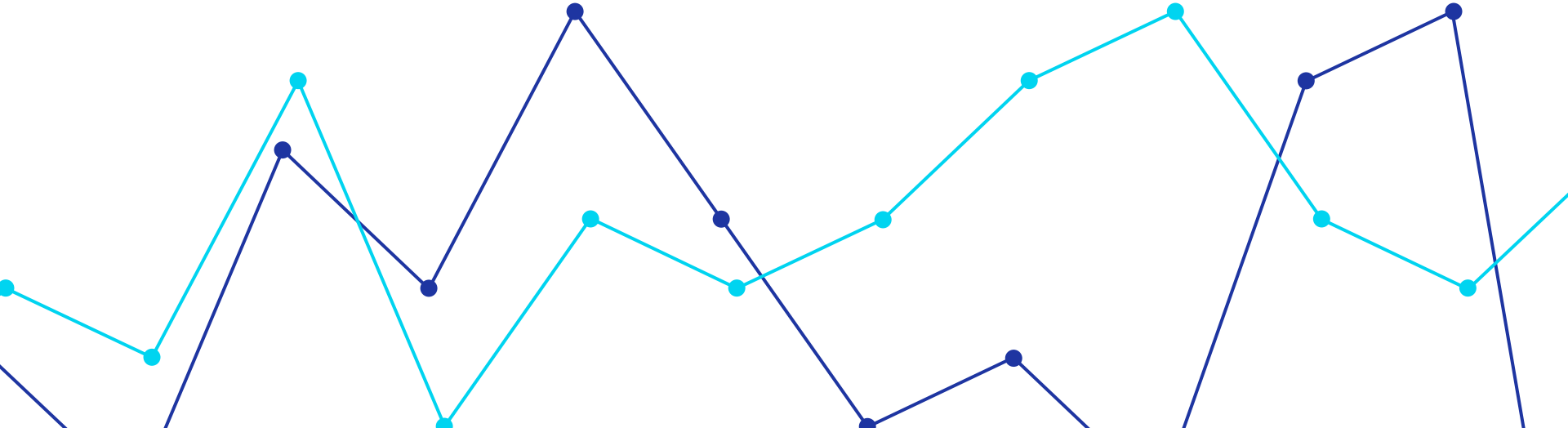


Trends in Median and Average Hourly US Wage Data from 1972 to 2022

Manav Bharath



Motivation

- Analyzing economic trends over time
- Changes in income inequality
- Policy implications
- Informed decisions from businesses, investors, and entrepreneurs of market opportunities and labor costs



Dataset

- Median and Average Hourly Wages in the USA (1973-2022)
 - Credit:
<https://www.kaggle.com/datasets/asaniczka/median-and-avg-hourly-wages-in-the-usa-1973-2022>
- Questions:
 - How do trends in median and average wages differ amongst various demographic groups?
 - How do the trends in average wages and median wages differ?
 - How do wage trends differ between different education levels?
 - Can we predict the average wages for the next 10 years?

Data Exploration

Wage Data Table

	year	median	average	men_median	men_average	women_median	women_average	white_median	white_average	black_median	...
0	2022	22.88	32.00	24.91	35.82	20.74	27.86	24.96	34.49	19.60	...
1	2021	23.05	32.08	24.93	35.61	21.46	28.28	25.40	34.50	19.45	...
2	2020	23.64	32.54	25.95	36.03	21.72	28.75	25.98	34.86	19.85	...
3	2019	22.12	30.36	24.04	33.65	20.42	26.83	24.39	32.79	18.45	...
4	2018	21.90	29.83	23.42	33.19	19.73	26.21	23.97	32.44	17.57	...
5	2017	21.55	29.21	23.78	32.27	19.69	25.91	23.87	31.64	17.84	...
6	2016	21.36	29.00	23.44	32.18	19.44	25.58	23.56	31.30	18.15	...
7	2015	20.99	28.43	23.19	31.52	19.07	25.11	23.37	30.75	17.38	...
8	2014	20.65	27.36	22.43	30.08	18.63	24.45	22.73	29.49	17.25	...
9	2013	20.63	27.54	22.63	30.49	18.87	24.38	22.73	29.72	17.61	...
10	2012	20.53	27.41	22.87	30.45	19.01	24.16	22.86	29.59	17.23	...

Features

year	int64	black_men_median	float64
median	float64	black_men_average	float64
average	float64	hispanic_men_median	float64
men_median	float64	hispanic_men_average	float64
men_average	float64	white_women_median	float64
women_median	float64	white_women_average	float64
women_average	float64	black_women_median	float64
white_median	float64	black_women_average	float64
white_average	float64	hispanic_women_median	float64
black_median	float64	hispanic_women_average	float64
black_average	float64	recent_high-school_graduate	float64
hispanic_median	float64	men_recent_high-school_graduate	float64
hispanic_average	float64	women_recent_high-school_graduate	float64
white_men_median	float64	recent_college_graduate	float64
white_men_average	float64	men_recent_college_graduate	float64
		women_recent_college_graduate	float64

Gender Wage Gap

- Added two new features:
gender_wage_gap_average and
gender_wage_gap_median
- Easier to analyze the gender wage gap
trends over time

	year	gender_wage_gap_average	gender_wage_gap_median
0	2022	7.96	4.17
1	2021	7.33	3.47
2	2020	7.28	4.23
3	2019	6.82	3.62
4	2018	6.98	3.69

Renamed Labels of Education Features

- Easier to write for the future when calling on the label

recent_high-school_graduate
men_recent_high-school_graduate
women_recent_high-school_graduate
recent_college_graduate
men_recent_college_graduate
women_recent_college_graduate



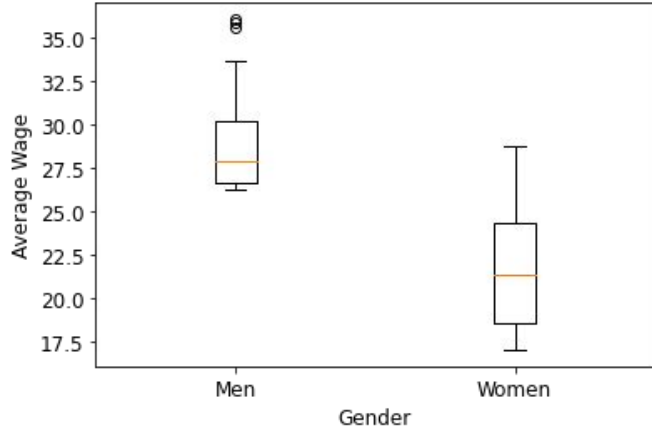
recent_HS_grad
men_recent_HS_grad
women_recent_HS_grad
recent_college_grad
men_recent_college_grad
women_recent_college_grad

Data Visualization

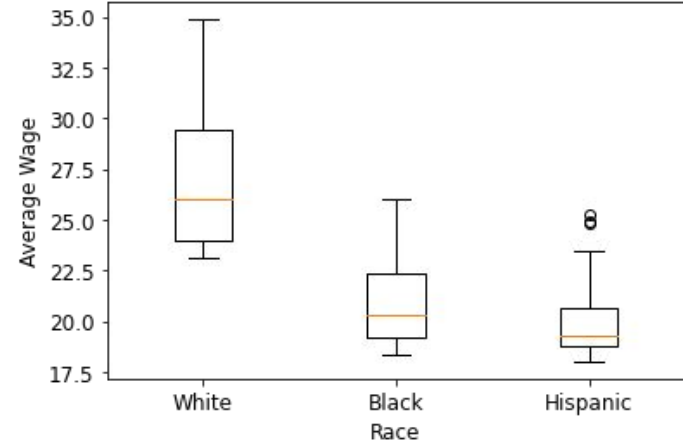
Demographic Impact on Wage Trends

Distribution of Average Wage

Distribution of Average Wage Across Gender from 1972 to 2022



Distribution of Average Wage Across Race from 1972 to 2022

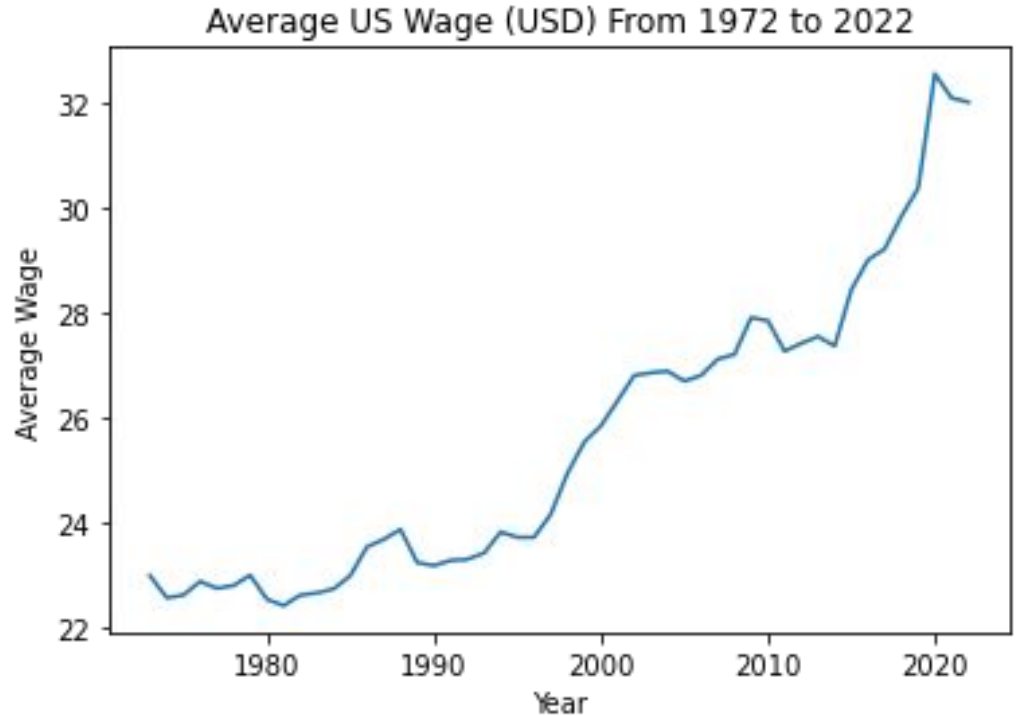


- Men have a higher wage distribution than women
- White people have a higher wage distribution than Black and Hispanic people, who have similar distributions

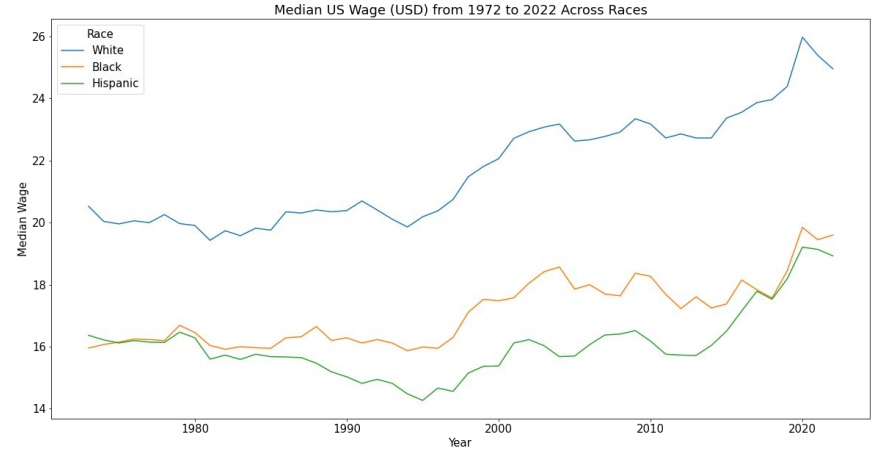
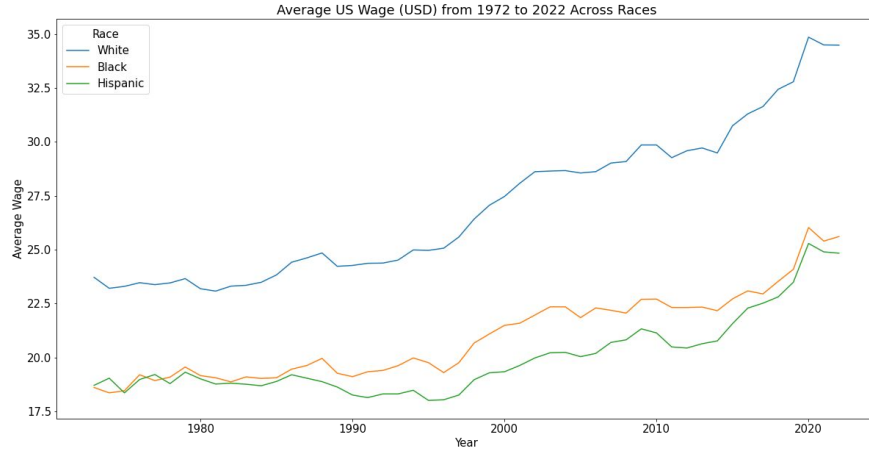
Average Wage Over Time

Key Insights

- Wage is increasing throughout time
- Occasional dips (90s, late 2000s, 2020)

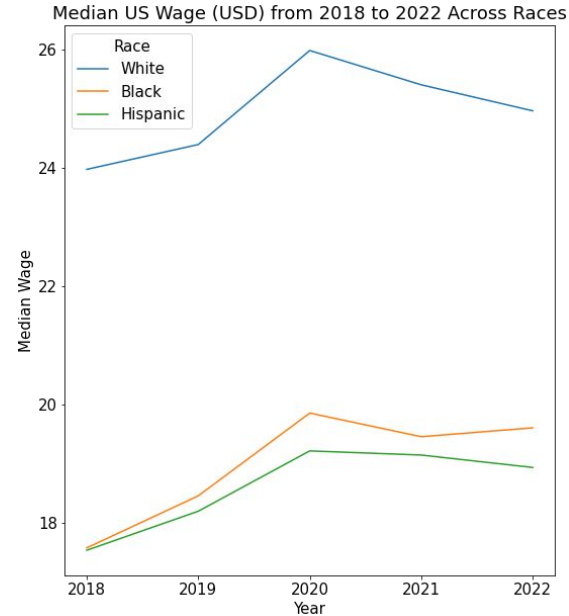
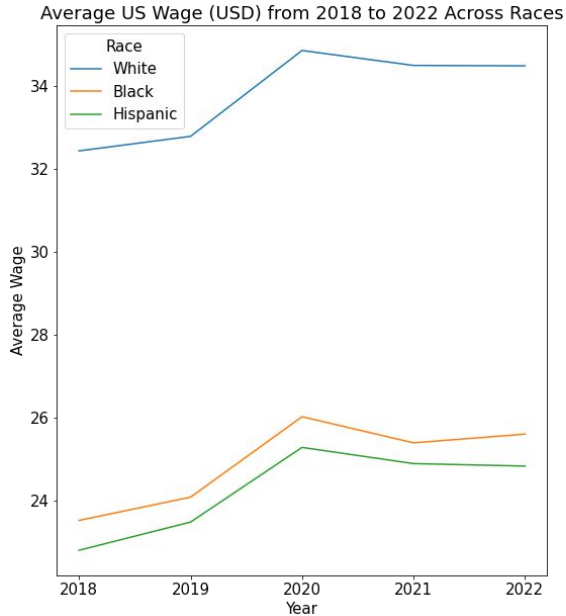


Average vs Median Wage Across Races



- Same general increase in average vs median wage
 - Decline in median wage seems to be sharper than declines in average wage
- Average and median wages of white people are significantly greater than for hispanic and black people

Average vs Median Wage Across Races

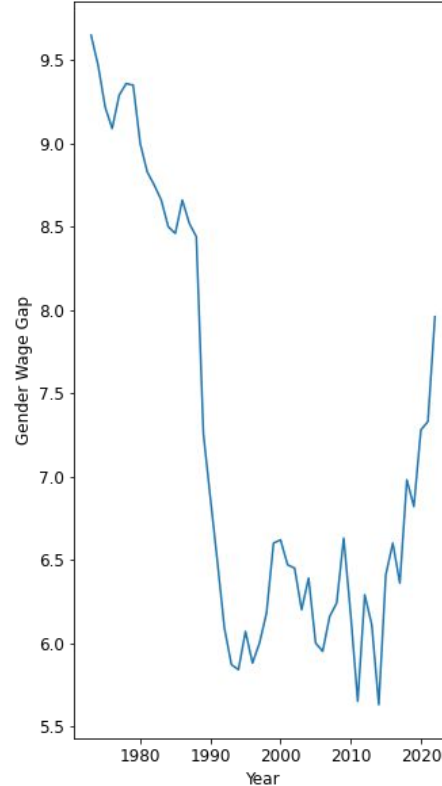


- Zoomed in for the past 5 years
- Average wage is around \$6-8 greater than median wage for ALL races

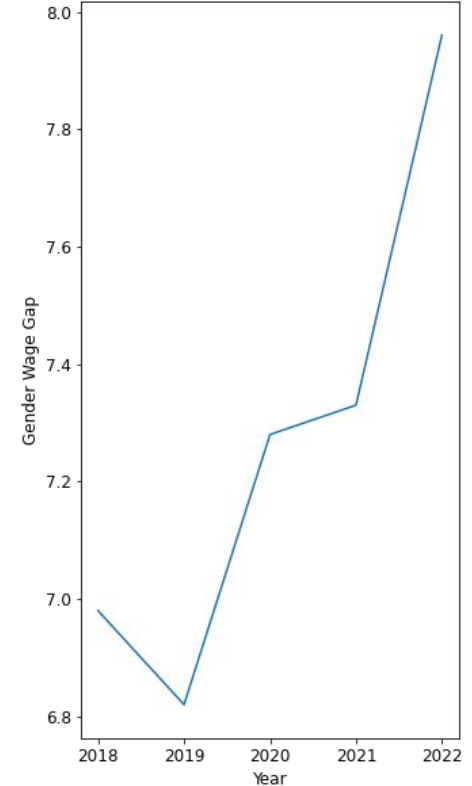
Gender Wage Gap Over Time

- The gender wage gap was significantly lower from 1990 to ~2012
- Loss of manufacturing jobs in the '90s (dominated by men) and the 2008 financial crisis left the average age of men and women similar to one another
- Recent rise in the average wage gap
 - Lower-paying jobs mostly consist of women

US Gender Wage Gap Average (USD) from 1972 to 2022

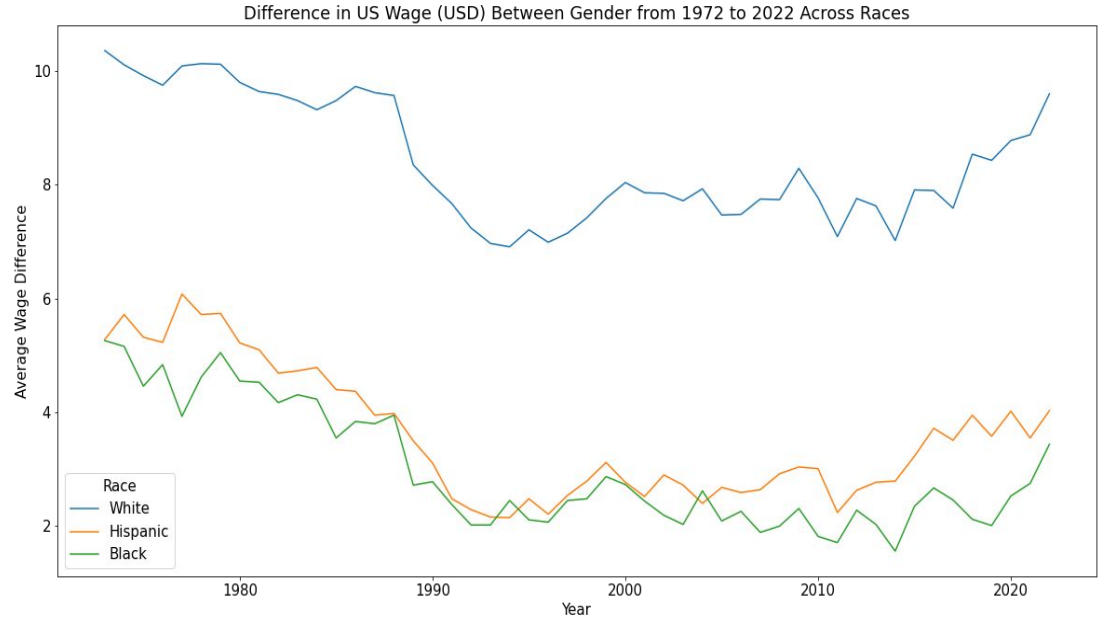


US Gender Wage Gap Average (USD) from 2018 to 2022



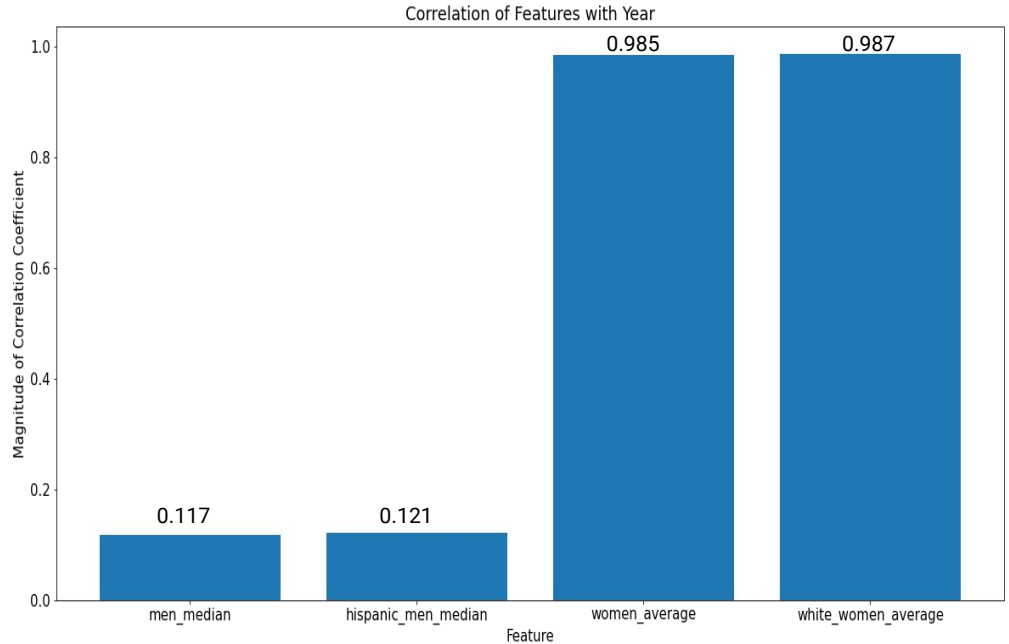
Gender Wage Gap Across Races

- Wage gap prevalent among all races
- Same general trend among all races
- Greater wage gap between white men and women than other races
 - Wider distribution in the wages of white people than other races

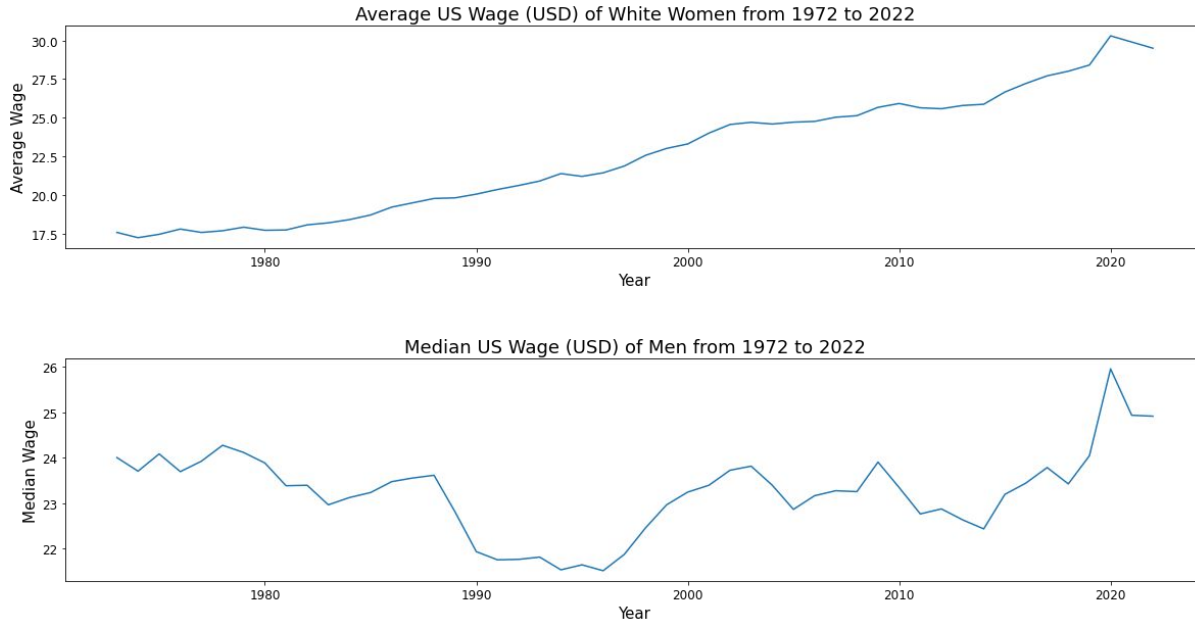


Correlation of Features with Year

- Select features represent the features with the lowest and highest correlation with year
- Correlation measured using the R-value (Pearson correlation coefficient)
- The growth rate of the average women's wage is highly linear
- The growth rate of men's wages is non-linear

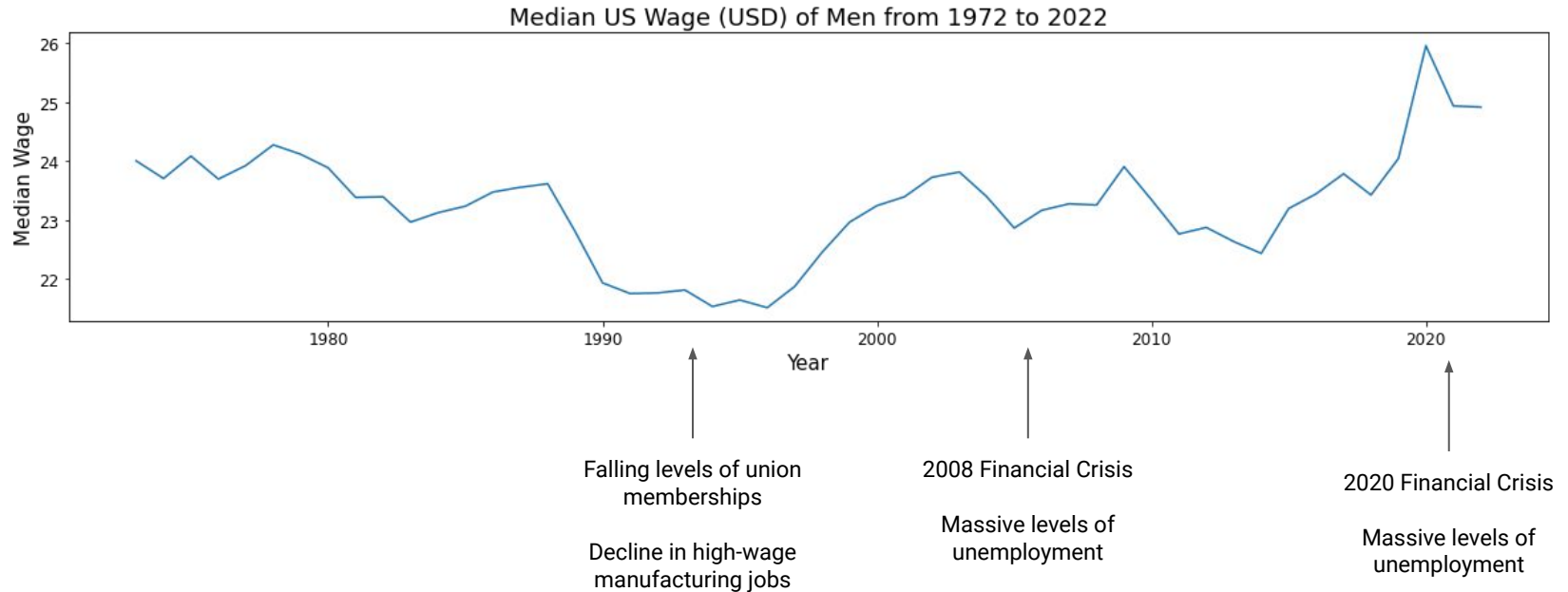


Median Male Wage vs Wage of White Women



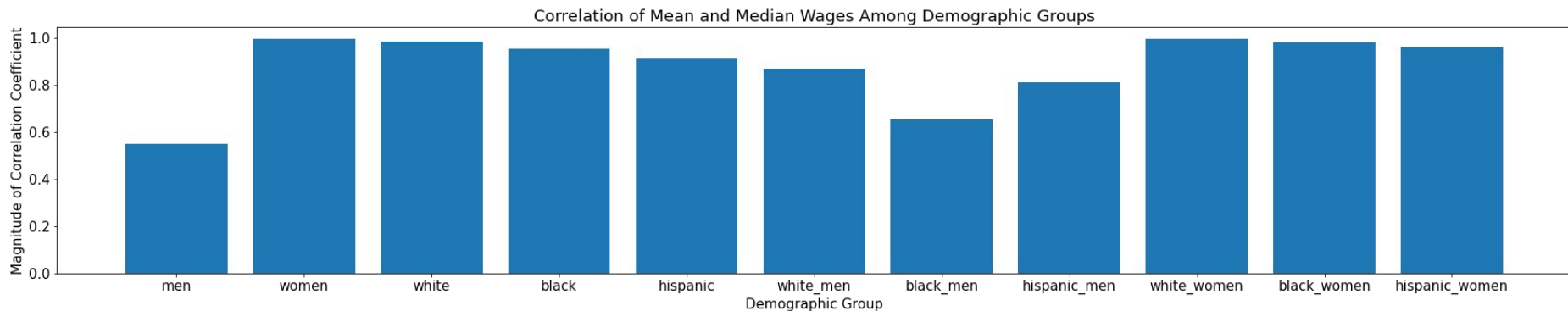
- Clear linear trend in white women's average wage and non-linear trend in median men's wage
- Growth rate of white women's wage is less impacted by socio-economic policy than male wages

Social/Economic Events Affect Wage



Mean/Median Comparisons

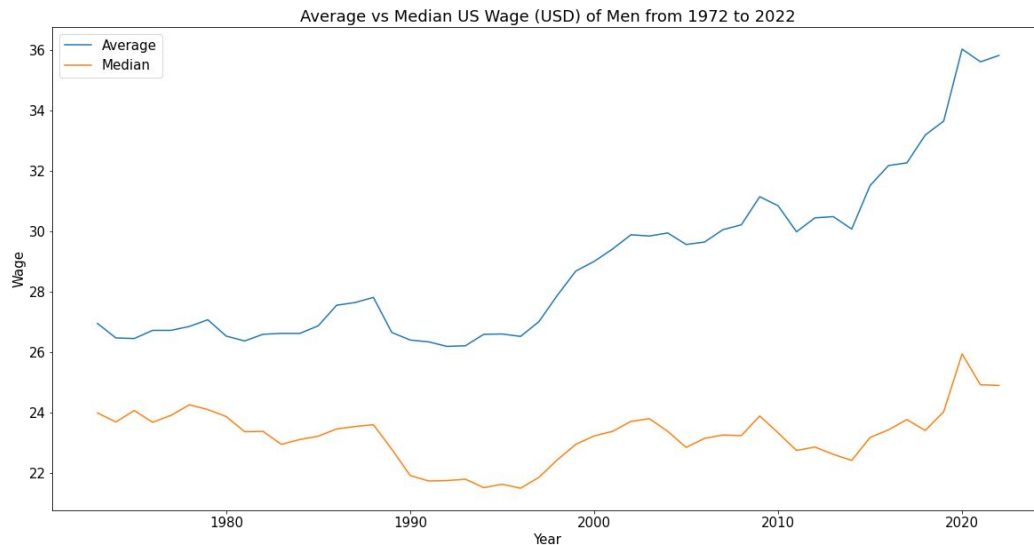
Mean/Median Correlations



- Mean and median are most correlated for women and least correlated for men across race
- Higher-earners (the 1%) mostly consist of men, so there is less correlation between male average wage and median wage

Mean Median Comparison in Wages of Men

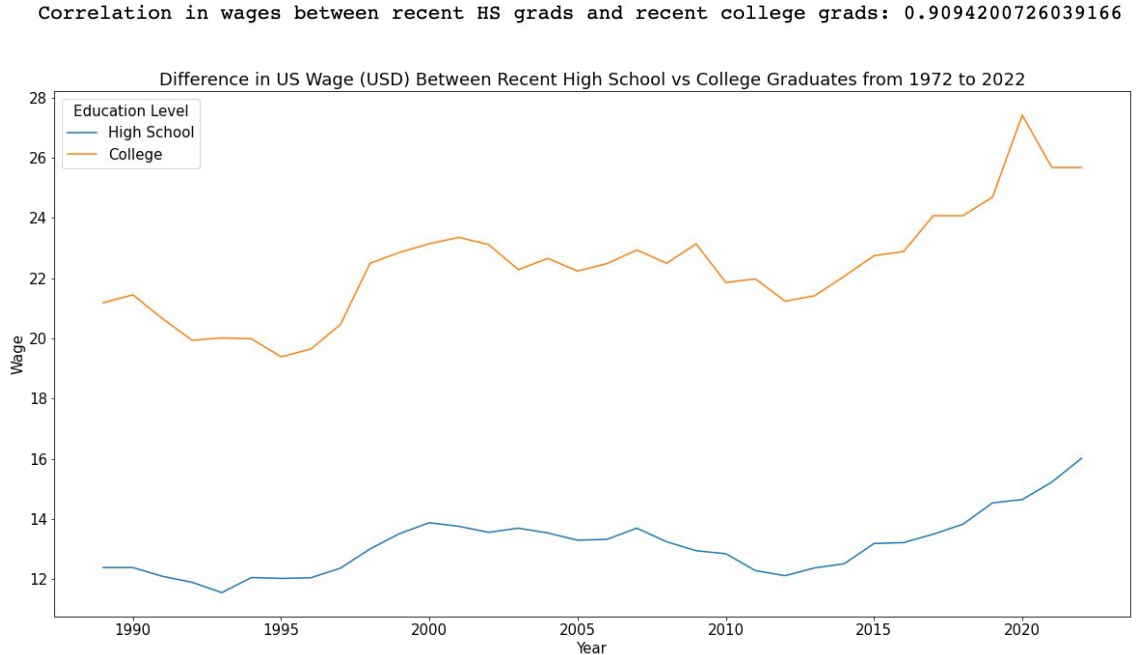
- Average wage is greater than median wage, caused by higher earners
- Sharper decrease in median wage in 2020 than average wage
- Increased wages of higher earners causes the mean to balance out more with drop in wages for rest of people



Impact of Education Level on Wage Trends

Recent High School Graduate Wage vs Recent College Graduate Wage

- Recent HS graduate wages are well correlated with recent college graduate wages
- Wages of recent college graduates are significantly greater than those of recent high school graduates
- The troughs and peaks of the wage trends for both high school and college align



Years of Lowest Wage for High School and College Grads

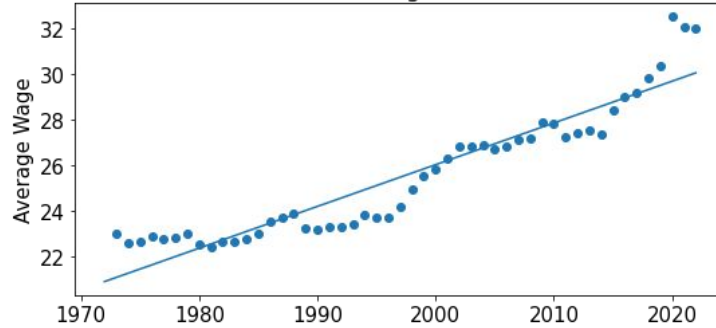
- The lowest wage-earning years for both high school and college education levels were the mid-90s
 - Loss in high-earning manufacturing jobs
 - Globalization
 - Falling union membership
- Early 2010s also include the lowest wage-earning years
 - Lasting effects from the 2008 financial crisis

	Years of Lowest Wage for Recent HS Grads	Years of Lowest Wage for Recent College Grads
0	1993	1995
1	1992	1996
2	1995	1992
3	1996	1994
4	1994	1993
5	1991	1997
6	2012	1991
7	2011	1989
8	1997	2012
9	2013	2013

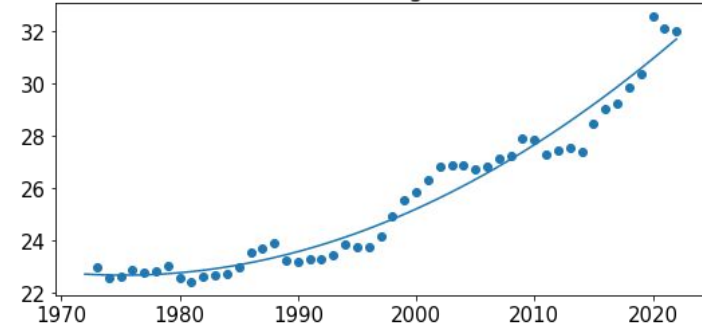
Modelling and Predictions

Using Different Models

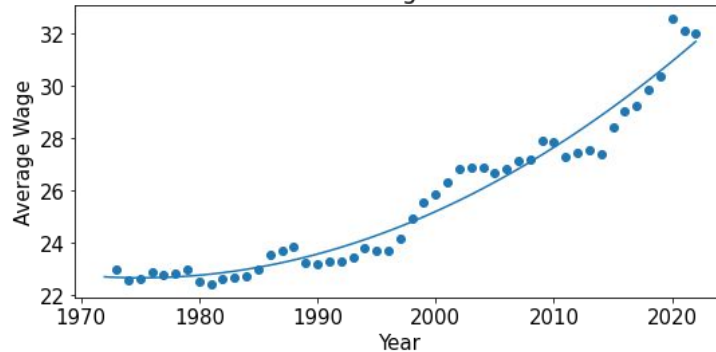
Linear Regression



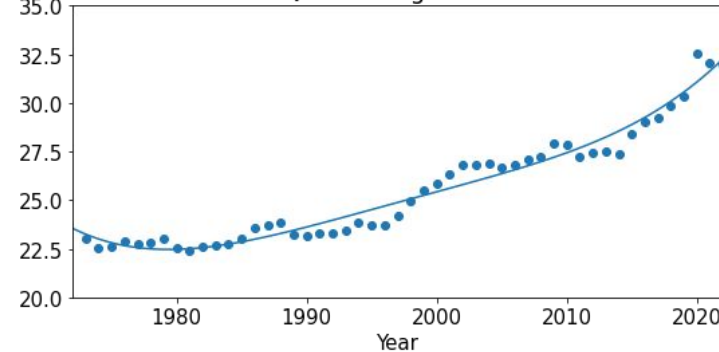
Quadratic Regression



Cubic Regression



Quartic Regression

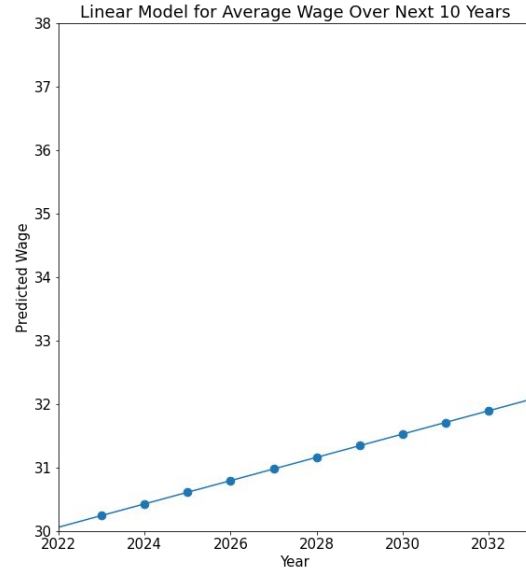
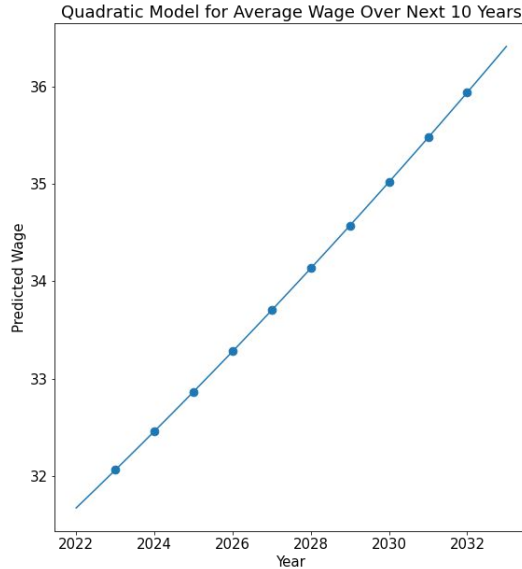


Using Different Models

Linear	R^2 : 0.8816487104036685
Quadratic	R^2 : 0.9554764973094224
Cubic	R^2 : 0.955477531241092
Quartic	R^2 : 0.9623740019749873

- Linear model has the lowest correlation
- Substantial increase in R^2 between linear and quadratic models (0.13)
- Increase as the degree of the polynomial model increases, but not substantial after quadratic
- Quadratic model is the best fit for the data

Modelling Wages Over the Next 10 Years



- Quadratic model predicts a much greater increase in wages than the linear model
- Linear model doesn't take into account the constant increase in growth rate

Additional Work

- I wanted to incorporate a classifier which would analyze wage data and determine the demographic data based on the wage trends, but I didn't have time to add it and wasn't quite sure how to pull it off.
- I wanted to see how well the quadratic model predicted the data by comparing it to future models of hourly wages from reputable sources, but I couldn't find that data online so I didn't pursue that analysis.