Data Analysis 2 Assignment 1

Jo kudo

Data

- Business Operations Specialists (500 - 740)
- Sample: 1634 males and 2288 females.

Distribution of Earnings

- Analysis of both normal and log distributions.
- Earnings gaps due to varying pay scales across different professions in the food industry.
- Log distribution approximates normal distribution, useful for further analysis.

Gender Gap with the Level of Education

Gender gap of logarithm wage is divided into three group based on "grade92", which means the highest educational grade completed.
- "grade92" = 32-36 (1) in the table below
    - Coeff is 0.0093. That is, earnings by women are 0.93% less than men.
    - $|t| = 0.068 < 1.96$, which means it is **not** statistically significant.
    - $p = 0.946 > 0.05$, which means it is **not** statistically significant.
- "grade92" = 36-41 (2)
    - Coeff is -0.1351. That is, earnings by women are 13.51% less than men.
    - $|t| = 4.514 > 1.96$, which means it is statistically significant.
    - $p = 0.000 < 0.05$, which means it is statistically significant.
- "grade92" = 42-46 (3)
    - Coeff is -0.2009. That is, earnings by women are 20.09% less than men.
    - $|t| = 9.220 > 1.96$, which means it is statistically significant.
    - $p = 0.000 < 0.05$, which means it is statistically significant.

Unconditional Earnings Analysis by Gender
- Overall, earnings by women are 18.7% less than men. (4)
- $|t| = 10.230 > 1.96$, which means it is statistically significant.
- $p = 0.000 < 0.05$, which means it is statistically significant.

Summary

- In terms of Business Operations Specialists, it seems the more grade they got, the wider the gender gap becomes.

	(1)	(2)	(3)	(4)
		Dependent variable: lnw		
female	0.009	-0.135***	-0.201***	-0.187***
	(0.137)	(0.030)	(0.022)	(0.018)
Constant	2.621***	3.072***	3.447***	3.330***
	(0.000)	(0.025)	(0.014)	(0.013)
Observations	13	1249	2660	3922
R^2	0.000	0.017	0.029	0.025
Adjusted R^2	-0.091	0.016	0.028	0.025
Residual Std. Error	0.417 (df=11)	0.502 (df=1247)	0.578 (df=2658)	0.577 (df=3920)
F Statistic	0.005 (df=1; 11)	20.376*** (df=1; 1247)	85.017*** (df=1; 2658)	104.657*** (df=1; 3920)
Note:				*p<0.1; **p<0.05; ***p<0.01