

# COGS108 Wi23

## Group032

### Final Presentation

Kendrick Nguyen, Tasnia Jamal, Ben Liou,  
Fanqi Lin





# Introduction

There has been numerous studies comprehending the relationship between personality and academic behaviors, and commonly, researchers use the Big Five personality traits as a model. The Big Five consists of openness, conscientiousness, extraversion, agreeableness, neuroticism. These personality factors are often described in the following by Verywell Mind blog [1].

- Openness refers to how willing a person is to exploring new experiences, thoughts, or interests. According to the blog, people who are high in openness tend to be more adventurous.
- Conscientiousness is a tendency to display self-discipline, act dutifully, and strive for achievement against measures or outside expectations. People high in conscientiousness tend to be more detail oriented.
- Extraversion is characterized by sociability and emotional expressiveness. People high in extraversion often gain energy from being around others.
- The agreeableness trait reflects individual differences in general concern for social harmony and how people-oriented a person is. Those with high agreeableness are expected to be more cooperative than competitive.
- Neuroticism refers to how likely a person is to react to negative situations. People who are high in neuroticism are often more prone to stress and anxiety.

[1] Cherry, Kendra. "What Are the Big 5 Personality Traits?" *Verywell Mind*, Verywell Mind, 11 Mar. 2023, <https://www.verywellmind.com/the-big-five-personality-dimensions-2795422>.



# Hypothesis

We hypothesize that conscientiousness and extraversion will strongly correlate to academic performance compared to the other Big Five personality traits. We assume having a disciplined and outgoing personality is a good outlook in the education setting, as suggested from our research.



# Dataset

## ***Dataset 1***

- Dataset Name: Engineering Graduate Salary Prediction
- Link to the dataset:  
<https://www.kaggle.com/datasets/manishkc06/engineering-graduate-salary-prediction>
- Number of observations: 2998 students
- Description: Originally used to predict the salary of an engineering graduate, the dataset provides labels, such as College GPA and AMCAT Personality test scores. The aforementioned labels will be used to measure academic performance and Big Five personality factors.

## ***Dataset 2***

- Dataset Name: Correlation of Personality Traits & GPA for Jordanian Medical Students
- Link to the dataset:  
<https://data.mendeley.com/datasets/5rwpwr9rf2/1>
- Number of observations: 307 students
- Description: A complete dataset on Hashemite University, Jordan's medical students, identifying their Big Five Model personality traits and College GPA. We plan to use this dataset in combination as it already yields the same labels as the first dataset.



# Cleaning

We broke down the data cleaning into two parts respectively for the Engineering Graduate Salary Prediction and Correlation of Personality Traits & GPA for Jordanian Medical Students datasets. The same process follows for both datasets where we:

- Created a dataframe by exclusively selecting our labels of interest, gender, gpa, openness, conscientiousness, extraversion, agreeableness, and neuroticism.
- Removed entries that contains any null (**NaN**) values
- Standardized the gender and gpa labels to m/f values and 0.0 - 4.0 scaling respectively
- Mapped each Big Five personality labels to percentiles relative to the dataset sample

Jordanian Medical Students datasets example,

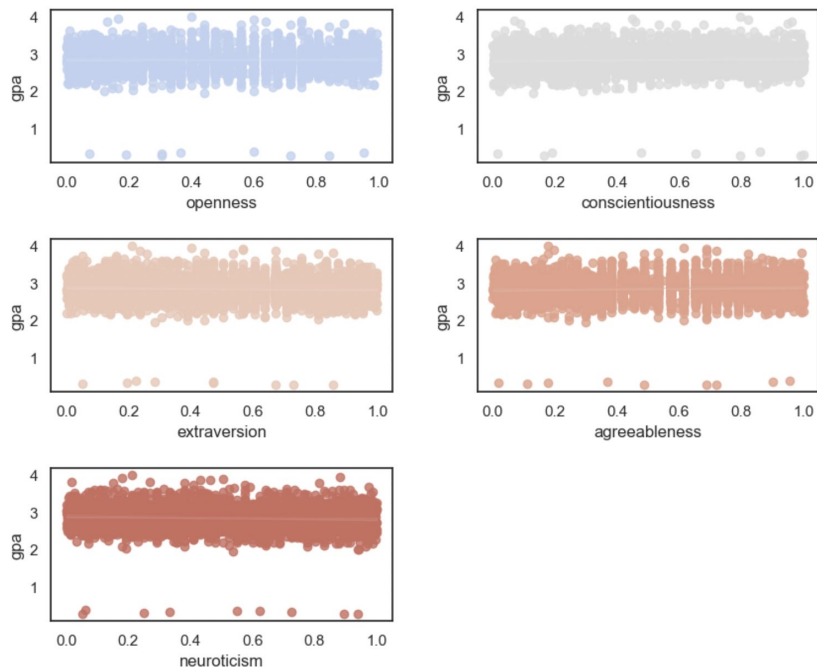
No.	Corresponding Year	Gender	Extraversion	Agreeableness	Conscientiousness	Neuroticism	Openness	GPA	Unnamed: 9	Gender	Unnamed: 11	Unnamed: 12	Unna
										1 = Male; Gender 2 = Female			
0	1	2	2	0.725	0.644	0.556	0.950	0.84	3.84	NaN	NaN	NaN	NaN
1	2	2	2	0.400	0.867	0.689	0.625	0.52	2.90	NaN	NaN	NaN	NaN
2	3	2	2	0.700	0.756	0.711	0.600	0.76	2.57	NaN	NaN	NaN	NaN
3	4	2	2	0.575	0.556	0.644	0.550	0.54	2.58	NaN	NaN	NaN	NaN
4	5	2	2	0.575	0.911	0.800	0.625	0.78	3.65	NaN	NaN	NaN	NaN



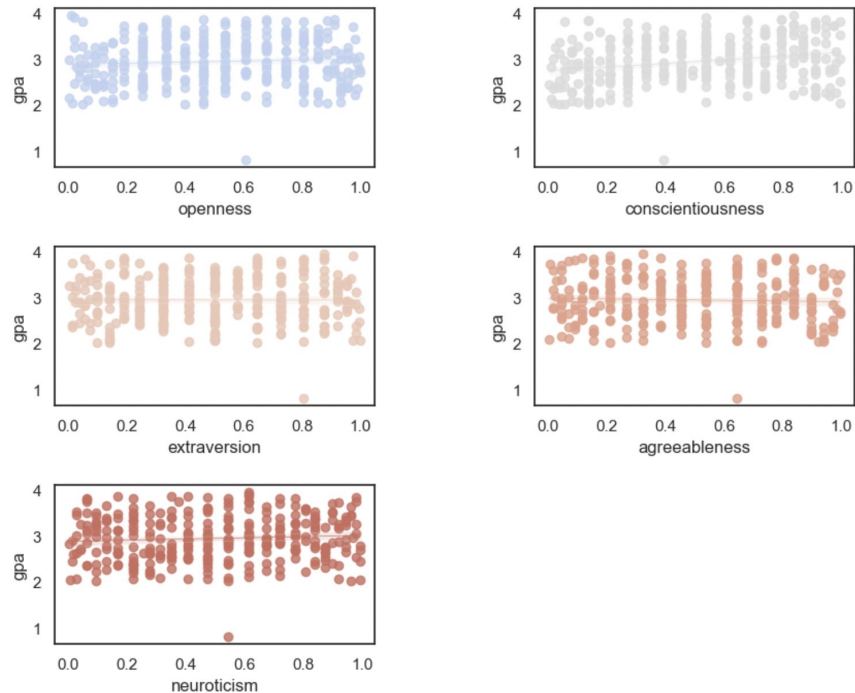
	gender	gpa	openness	conscientiousness	extraversion	agreeableness	neuroticism
0	f	3.84	0.887622	0.136808	0.806189	0.153094	0.980456
1	f	2.90	0.057003	0.452769	0.035831	0.838762	0.543974
2	f	2.57	0.677524	0.537459	0.726384	0.452769	0.472313
3	f	2.58	0.068404	0.333876	0.324104	0.043974	0.350163
4	f	3.65	0.744300	0.793160	0.324104	0.941368	0.543974

# EDA (single-variable analysis)

Engineering Graduate Dataset Scatterplots



Medical Students Dataset Scatterplots





# EDA (multi-variable analysis)

## Interpretation of OLS Regression: Engineering Students

	coef	std err	t	P> t	[0.025	0.975]
<b>Intercept</b>	2.9917	0.024	123.722	0.000	2.944	3.039
<b>openness</b>	-0.0196	0.024	-0.822	0.411	-0.066	0.027
<b>conscientiousness</b>	0.0221	0.024	0.917	0.359	-0.025	0.069
<b>extraversion</b>	-0.1016	0.023	-4.407	0.000	-0.147	-0.056
<b>agreeableness</b>	0.0869	0.025	3.425	0.001	0.037	0.137
<b>neuroticism</b>	-0.0605	0.022	-2.796	0.005	-0.103	-0.018
<b>gender</b>	-0.1247	0.014	-9.073	0.000	-0.152	-0.098

Here we can conclude that extraversion, agreeableness, and neuroticism personalities have coefficients of -0.1016, 0.0869, and -0.0605 respectively. Or in other words, strictly the extraversion,, agreeableness, neuroticism personalities are not correlated to gpa.

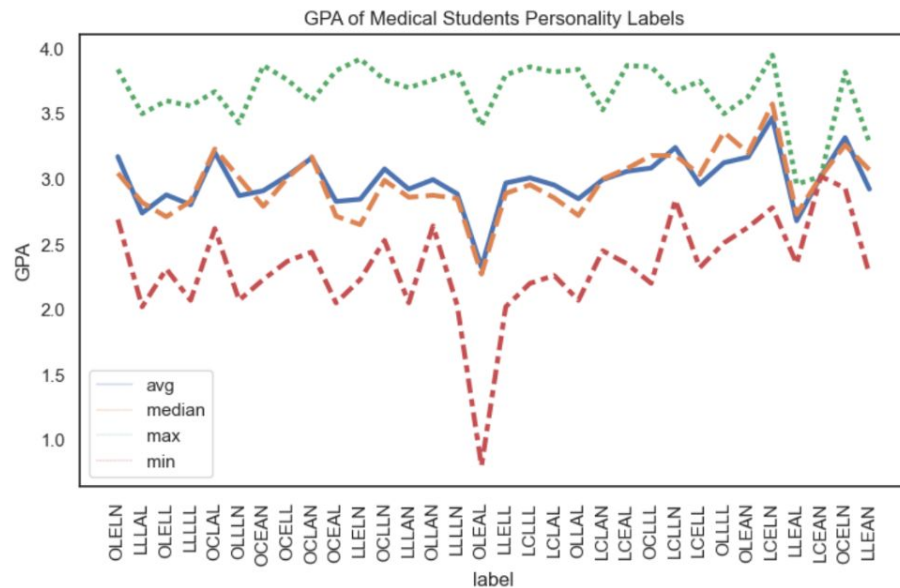
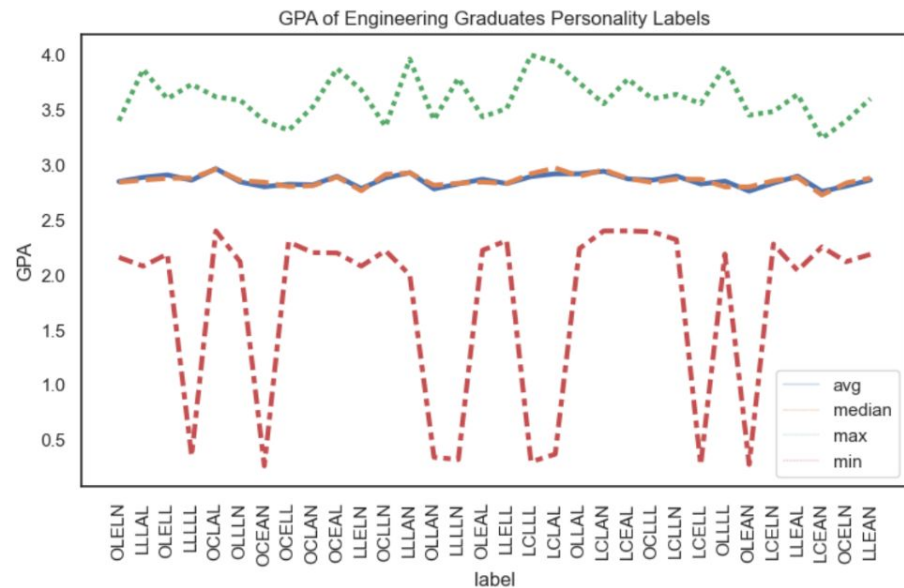
## Interpretation of OLS Regression: Medical Students

	coef	std err	t	P> t	[0.025	0.975]
<b>Intercept</b>	2.7733	0.117	23.733	0.000	2.543	3.003
<b>openness</b>	-0.0095	0.102	-0.093	0.926	-0.211	0.192
<b>conscientiousness</b>	0.5206	0.106	4.927	0.000	0.313	0.729
<b>extraversion</b>	-0.0590	0.102	-0.579	0.563	-0.260	0.141
<b>agreeableness</b>	-0.1948	0.100	-1.941	0.053	-0.392	0.003
<b>neuroticism</b>	0.1779	0.106	1.686	0.093	-0.030	0.386
<b>gender</b>	-0.0861	0.057	-1.521	0.129	-0.198	0.025

Here we can conclude the conscientiousness personality has a coefficient of 0.5206. Or in other words, strictly conscientiousness is moderately correlated to gpa.



## EDA (ii)



We created a labeling convention for a Big Five type. The initial of each trait represents a "high" score and the letter "L" represents a "low" score. For example, OCLLN would mean a high openness, conscientiousness, neuroticism and score, but a low extraversion and agreeableness score.





# Conclusion

The analysis we performed suggested that the field of study may be a confounding variable; therefore, the conclusions we draw are conditional to a specific dataset/student. We observed that strictly the conscientiousness, extraversion, agreeableness, and neuroticism traits have either no or moderate correlation to GPA. Specifically to the (dataset) population of engineering students, the extraversion, agreeableness, and neuroticism traits have zero/no correlation to GPA. Likewise, specifically to the (dataset) population of medical students, the conscientiousness trait is moderately and positively correlated to GPA.

We were also able to conclude that specific Big Five types have a slightly higher correlation to academic performance than others.

For both datasets, we observed that there is generally no trend among most of the personality types that we had assigned. Although again, for medical, there were two sharp peaks - OLEAL performed the worst while LCELN did the best. Moreover, a few other local peaks in the medical plot were low in agreeableness but high in neuroticism.



# Ethics and Privacy

- Found publicly on Kaggle
- Jordanian dataset was anonymous, but Graduate dataset had student ID's
- Potentially matchable with names linked on publicly posted resumes
- Companies may use to better understand job applicants, or discriminate (dual nature)



# Limitations and Moving Forward

- Can't generalize to all students:
  - only engineering and medical students
  - only Indian and Jordanian students
  - college students
- Maybe we can consider a more wide ranging measure of "success" for our dependent variable, besides GPA
  - Wealth: would apply to individuals outside of school, but not children
  - IQ test score: would apply to all ages
  - Life Happiness Score: would apply to all ages



# References

Rosito, Asina Christina. "The Big Five Personality Traits, Self-regulated Learning, and Academic Achievement." Retrieved 5 March 2023, from <https://www.scitepress.org/Papers/2018/85910/85910.pdf>

Gandhi, Vandana. "Big Five Personality Traits and Academic Performance of Post Graduate Students." Retrieved 5 March 2023, from [https://www.researchgate.net/publication/329446960\\_Big\\_Five\\_Personality\\_Traits\\_and\\_Academic\\_Performance\\_of\\_Post\\_Graduate\\_Students](https://www.researchgate.net/publication/329446960_Big_Five_Personality_Traits_and_Academic_Performance_of_Post_Graduate_Students)