Big Data Analytics :

Kaggle Project : Exploring mental health data

Hamida AHMED ALI u32003122

Samia BASAID u21805466

Jina EL HABER u32316105

Mohamed Amine TALEB u32414583

MASTER 2 MCI GIT

I / Background and Purpose of the Project:Exploring Mental Health Data

Around 41% of university students are experiencing depressive symptoms, a significant increase from the 26% reported before the COVID-19 pandemic according to the University of Bordeaux.

Kaggle projects are widely recognized and respected in the data science community

II / Methodology and Scope of the Project

id	Name	Gender	Age	City
0	Aaradhya	Female	49.0	Ludhiana
1	Vivan	Male	26.0	Varanasi
2	Yuvraj	Male	33.0	Visakhapatnam
3	Yuvraj	Male	22.0	Mumbai
4	Rhea	Female	30.0	Kanpur
5	Vani	Female	59.0	Ahmedabad
6	Ritvik	Male	47.0	Thane
7	Rajveer	Male	38.0	Nashik

submission.csv (844.21 kB)

id	class
140700	0
140701	0
140702	0
140703	1
140704	0
140705	0
140706	0
140707	0

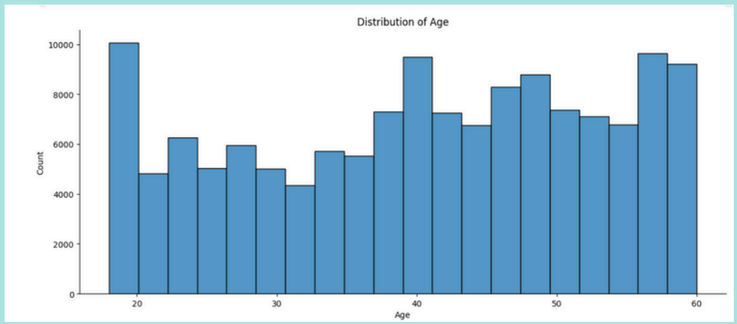
THE MAIN STEPS OF THE NOTEBOOK

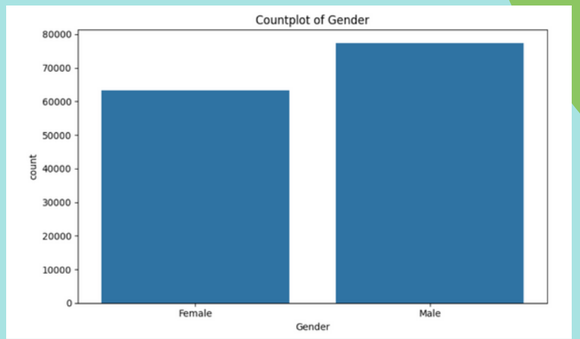
- Import necessary librairies
- Reading and Understanding our Data
- Exploratory Data Analysis
- Feature Engineering and Data Preprocessing
- Model Training

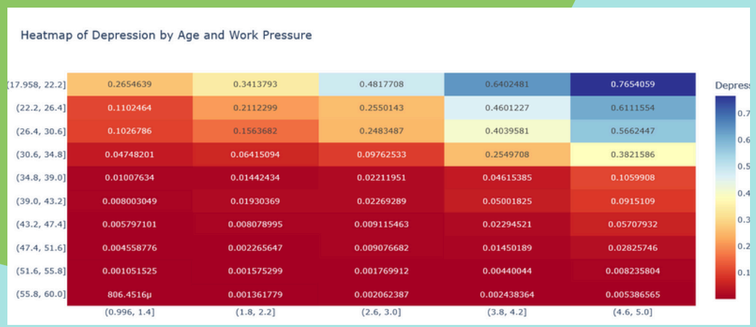
Preparing the data in this way was essential to ensure the success of our Kaggle project and help our machine learning model to explore depression levels accurately.

III. Code Analysis: Visualizations and Outputs

- **This project created a lot of output based on visualization**
- **The Exploratory Data Analysis (EDA) phase**
- **These are examples of graphs we got from the code**





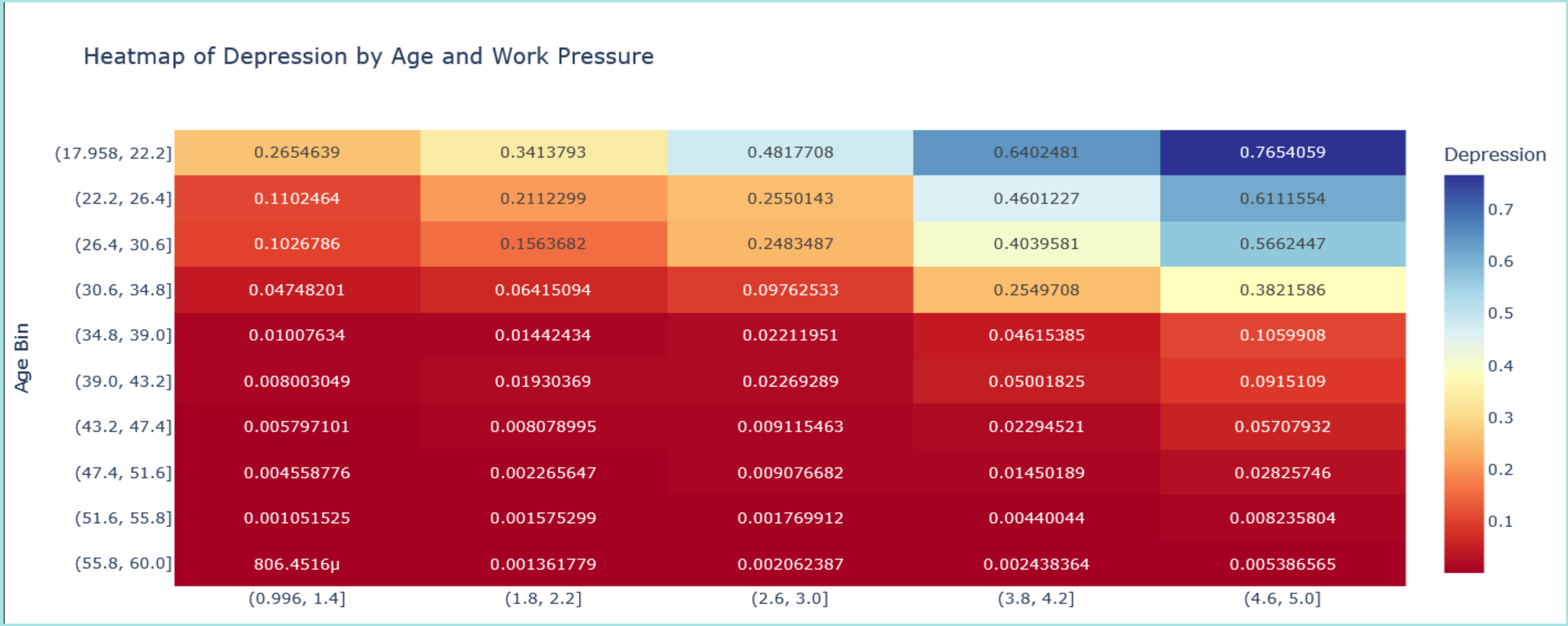


HISTOGRAMS

BAR CHARTS

HEATMAPS

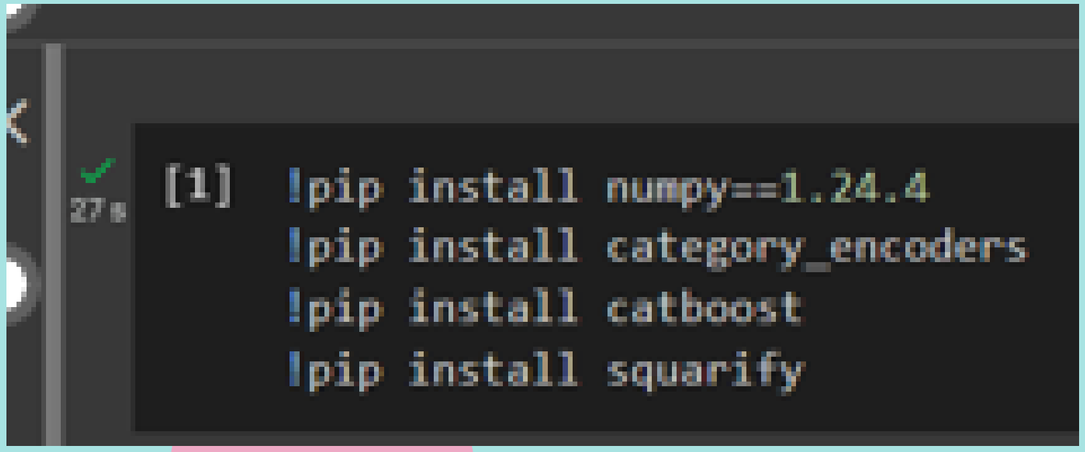
Interpretation of this heatmap



IV. Issues Encountered and how we solved them

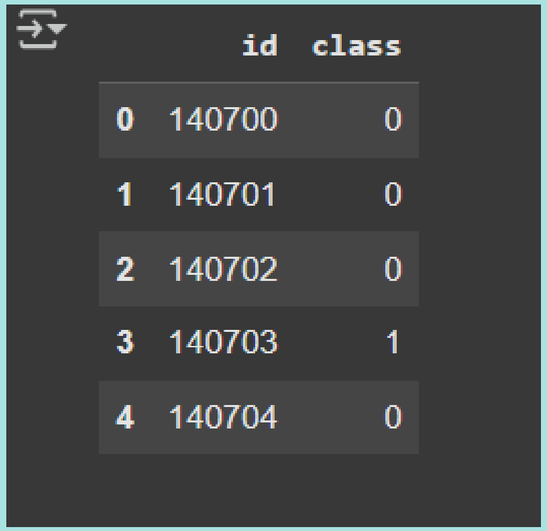
The main issues that we went through creating the code on Google Collab were** library imports errors **and** running the code at the same time **as other people of our groupe :

- *Not having the librairies installed by default in Google Collab (example : 'squarify'*
- *Solution :*



V. Recommendations on the code and the project

- A notebook with more detailed explanations
- Change the format of the output at the end of the notebook so that we can make better interpretations ?



	id	class
0	140700	0
1	140701	0
2	140702	0
3	140703	1
4	140704	0

CONCLUSION

- We successfully reimplemented a complete Kaggle notebook within the Google Colab environment.
- We adapted the original code to function seamlessly within our Colab environment.
- Most importantly, this project enabled us to engage with a socially relevant topic that deeply resonates with our values: mental health.