


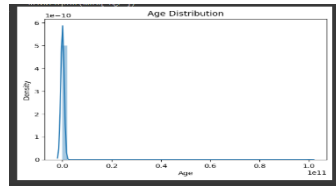
## Data Collection and Preprocessing Phase

Date	15 June 2024
Team ID	739903
Project Title	Mental health prediction
Maximum Marks	6 Marks

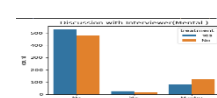
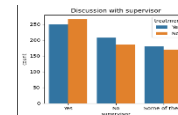
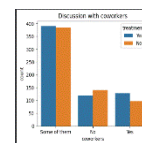
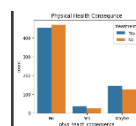
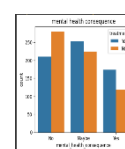
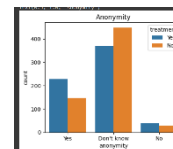
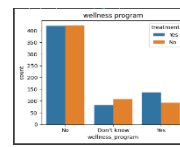
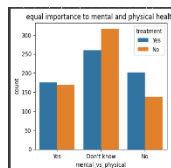
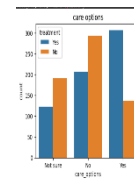
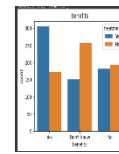
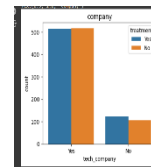
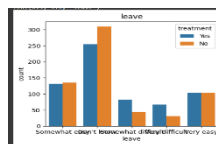
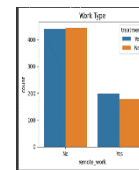
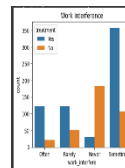
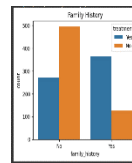
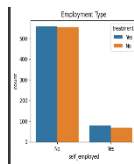
### Data Exploration and Preprocessing Report

Data exploration for mental health prediction involves analyzing and understanding datasets to identify patterns and correlations relevant to mental health factors. Preprocessing includes cleaning data, handling missing values, and transforming variables to prepare for predictive modeling and analysis.

Section	Description
Data Overview	<p><u>Dimension:</u> 5 rows*27 columns</p>  <p><u>Descriptive statistics:</u></p>
Univariate Analysis	



## Bivariate Analysis



## Multivariate Analysis

## Outliers and Anomalies

-

## Data Preprocessing Code Screenshots

### Loading Data

```
[ ] data = pd.read_csv("../content/survey.csv")
```

```
data.head()
```

	Timestamp	Age	Gender	Country	state	self_employed	family_history	treatment	work_interfere	no_employees	...	leave	mental_health_consequence	phys_hoa
0	2014-08-11:29:51	27	Female	United States	IL	NaN	No	Yes	Often	6-25	...	Somewhat easy	No	
1	2014-08-11:29:57	44	M	United States	IN	NaN	No	No	Rarely	More than 1000	...	Don't know	Maybe	
2	2014-08-11:29:44	32	Male	Canada	NaN	NaN	No	No	Rarely	6-25	...	Somewhat difficult	No	
3	2014-08-11:29:46	31	Male	United Kingdom	NaN	NaN	Yes	Yes	Often	26-100	...	Somewhat difficult	Yes	
4	2014-08-11:20:22	31	Male	United States	TX	NaN	No	No	Never	100-500	...	Don't know	No	

5 rows x 27 columns

### Handling Missing Data

```
data.isnull().sum()
```

Timestamp	0
Age	0
Gender	0
Country	0
state	515
self_employed	10
family_history	0
treatment	0
work_interfere	264
no_employees	0
remote_work	0
tech_company	0
benefits	0
care_options	0
wellness_program	0
seek_help	0
anonymity	0
leave	0
mental_health_consequence	0
phys_health_consequence	0
coworkers	0
supervisor	0
mental_health_interview	0
phys_health_interview	0
mental_vs_physical	0
obs_consequence	0
comments	1095
dtype: int64	

Data Transformation	
Feature Engineering	Attached the codes in final submission.
Save Processed Data	-