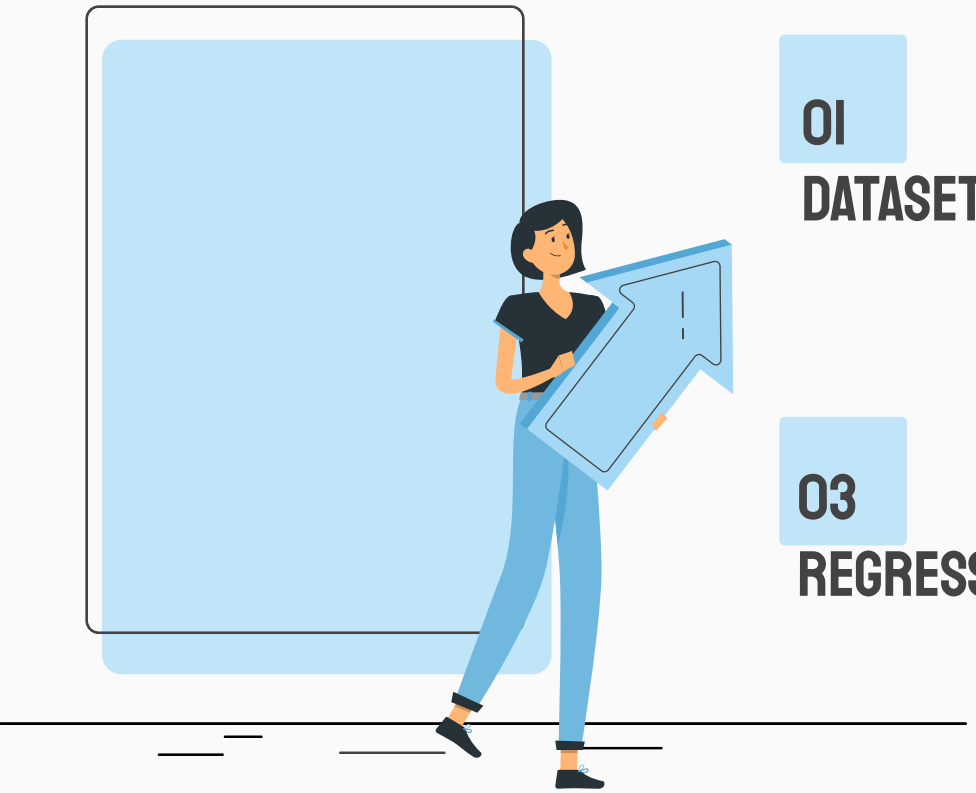
SOCIAL MEDIA'S IMPACT ON MENTAL HEALTH

Group 2: Margie Jurado Espinoza, Vrinda Arora, Mohammed Buharie, Phuong Trang Tran



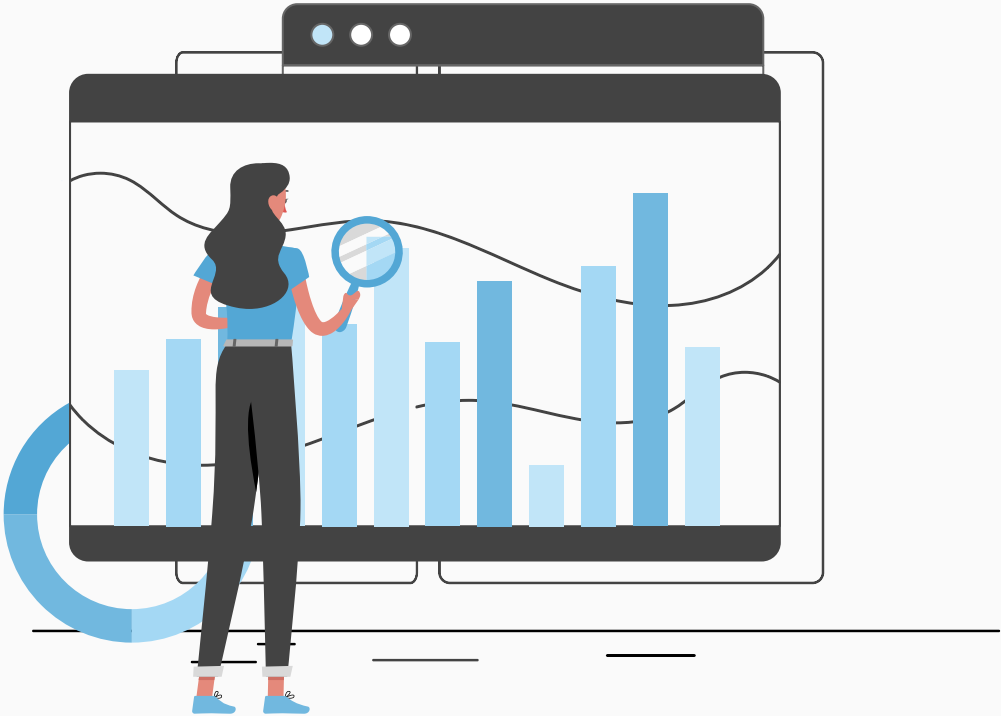
01
DATASET

02
ASSOCIATION ANALYSIS

03
REGRESSION ANALYSIS

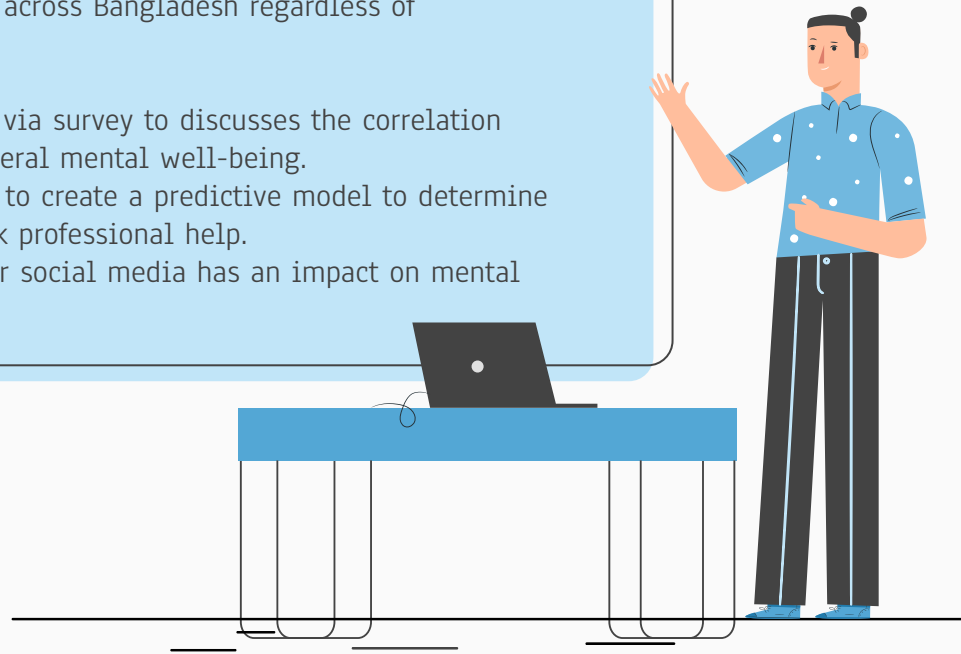
04
FINAL ANALYSES

01.
DATASET



THE DATASETS

- **Dataset 1:**
  - Unique responses from 791 adults via survey collected in 2021
  - Relates to the four dimensions of psychological distress including depression, anxiety, loneliness, and sleep disturbances.
  - Responses were collected from all across Bangladesh regardless of sociodemographic background.
- **Dataset 2:**
  - 481 responses from various groups via survey to discusses the correlation between social media use and general mental well-being.
  - Used machine learning techniques to create a predictive model to determine whether the individual should seek professional help.
- Both datasets will help determine whether social media has an impact on mental health using RStudio for data analysis.



DATA CLEANING

01
REMOVAL OF UNNECESSARY DATA

Columns such as specific sleep habits that were not needed for our analysis were removed.

02
RENAMING OF COLUMNS

Altered the name of columns that were unnecessarily long/complicated.

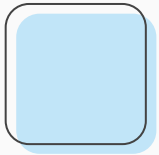
03
REPLACED MISSING VALUES

Some columns had missing data which was replaced with N/As.

04
TRANSFORMED VARIABLES

Converted some fields into Factor vectors so we could create tables and plots.

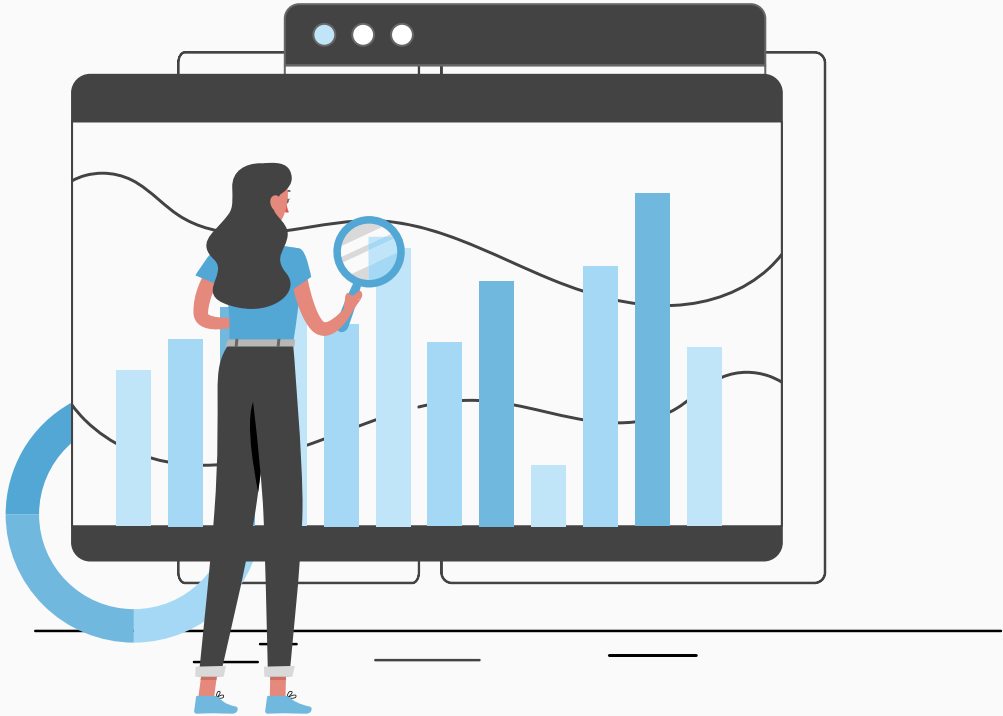
OTHER METHODS USED



DPLYR PACKAGE

The dplyr package helps with data manipulation and provides a set of verbs that help you solve the most common data manipulation challenges such as mutate, filter, summarise, and arrange.

02.
ASSOCIATION ANALYSIS



ASSOCIATION ANALYSIS

VARIABLES:

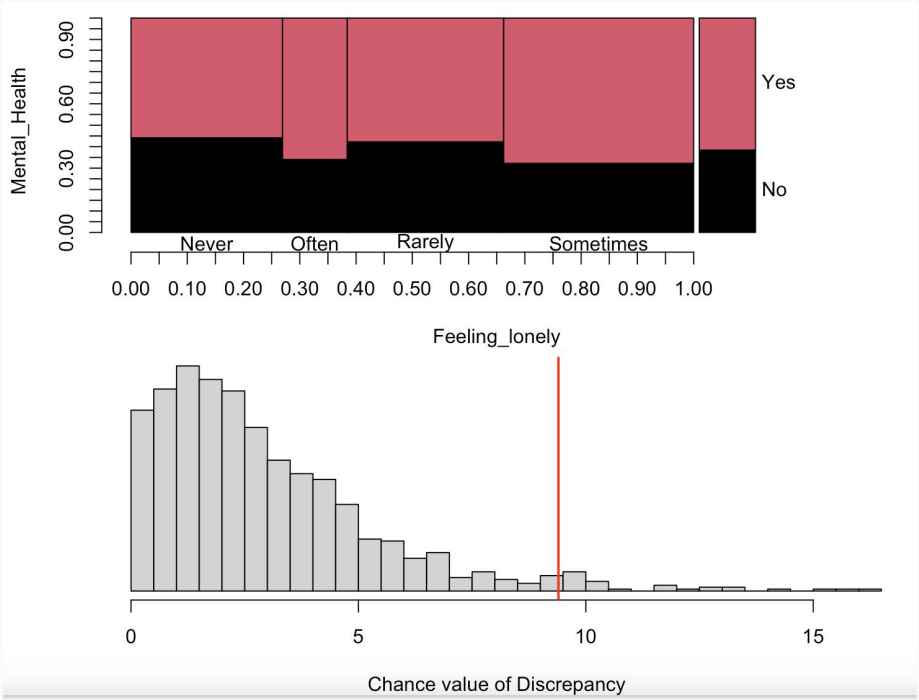
"Do you think your mental wellbeing would be better if you do not use social media?"

&

"In the past 30 days, do you feel a lack of companionship?"

Conclusion: P-value is less than 0.05 so there exists an association.

```
Permutation procedure:
  Discrepancy Estimated p-value
      9.39752             0.029
With 1000 permutations, we are 95% confident that:
 the p-value is between 0.02 and 0.041
If 0.05 is in this range, change permutations= to a larger number
```



ASSOCIATION ANALYSIS

VARIABLES:

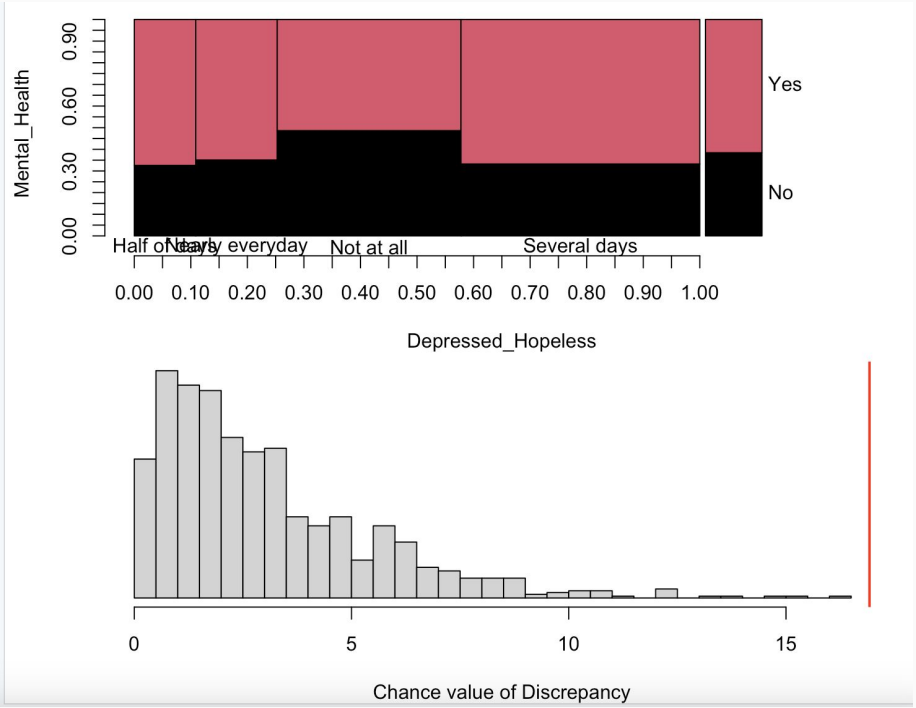
"Do you think your mental wellbeing would be better if you do not use social media?"

&

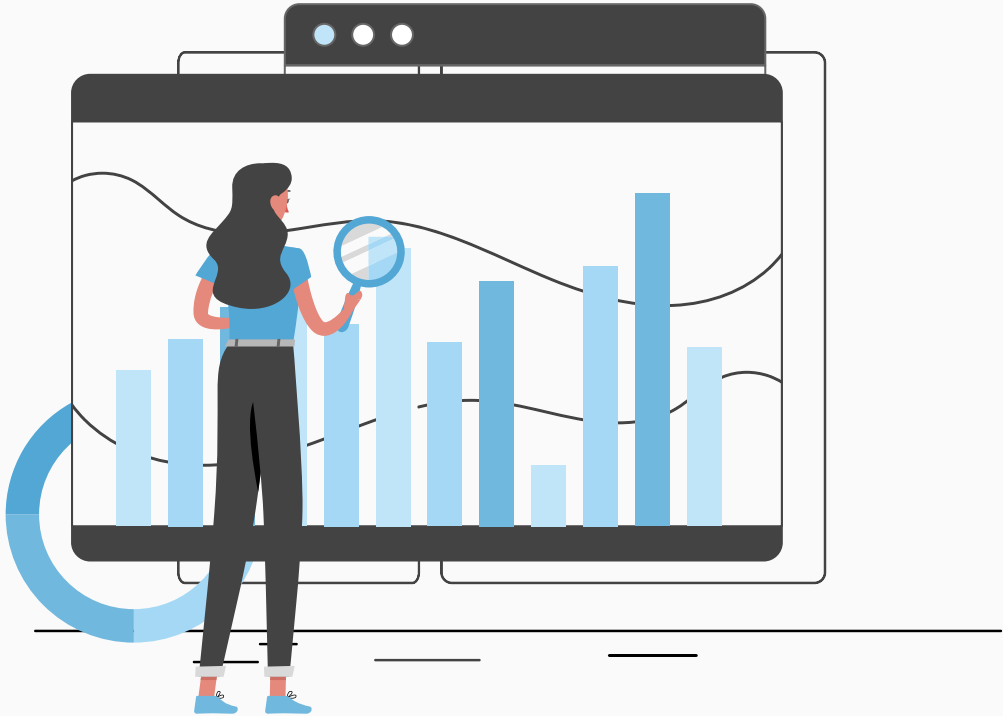
"In the last 30 days: feeling down, depressed or hopeless?"

Conclusion: P-value is less than 0.05 so there exists an association.

```
Permutation procedure:
  Discrepancy Estimated p-value
     16.92116                 0
With 1000 permutations, we are 95% confident that:
 the p-value is between 0 and 0.004
If 0.05 is in this range, change permutations= to a larger number
```



03.
REGRESSION ANALYSIS



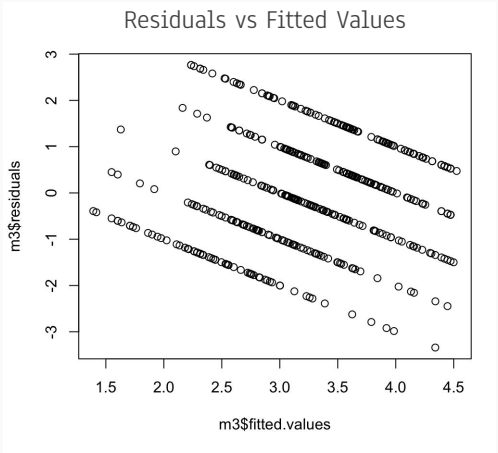
Model 3

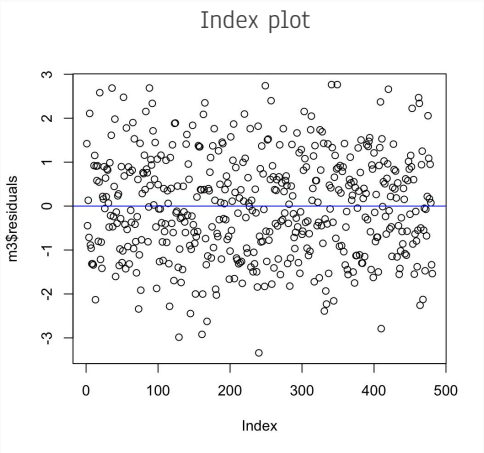
REGRESSION ANALYSIS

VARIABLES:

Concentration compared to comparison, sleep quality, and age
Residual vs Fitted Values plot shows **homoscedasticity**

Summary:
Residual standard error is 1.165 percentage points when predicting the response rate

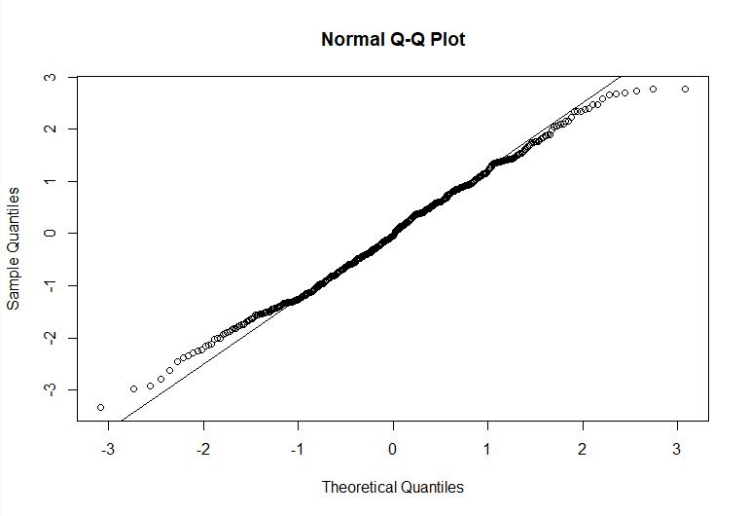





```
Analysis of Variance Table

Response: Concentration
               Df Sum Sq Mean Sq F value    Pr(>F)
Comparison      1 119.99 119.993  88.480 < 2.2e-16 ***
Sleep_quality   1  72.37  72.367  53.362 1.173e-12 ***
Age             1  31.80  31.803  23.451 1.735e-06 ***
Residuals     477 646.89   1.356
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

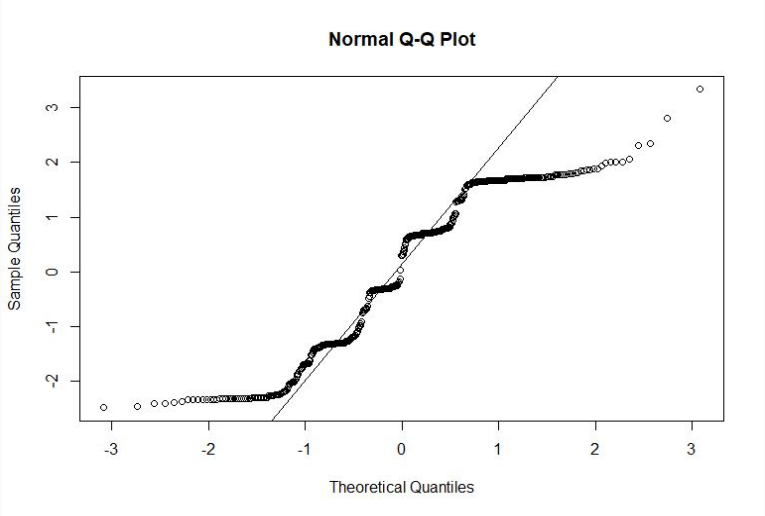
```
Residuals:
    Min      1Q  Median      3Q     Max
-3.3408 -0.8470 -0.0465  0.8426  2.7651

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)    2.334385   0.227092  10.279  < 2e-16 ***
Comparison     0.288902   0.038562   7.492 3.32e-13 ***
Sleep_quality  0.244173   0.037474   6.516 1.84e-10 ***
Age           -0.026357   0.005443  -4.843 1.73e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.165 on 477 degrees of freedom
Multiple R-squared:  0.2573,    Adjusted R-squared:  0.2527
F-statistic:  55.1 on 3 and 477 DF,  p-value: < 2.2e-16
```

INTERPRETATION OF Q-Q PLOT



Q-Q PLOT FOR MODEL 3:

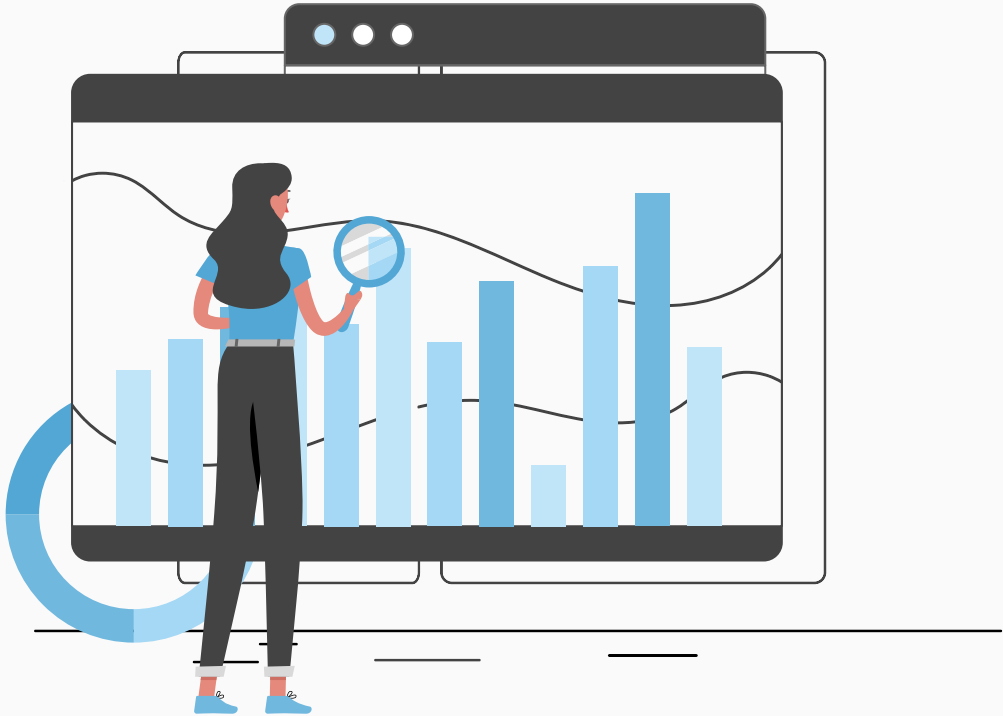
- The observation in the plot seem to follow a linear pattern..
- No significant deviation from the regression line.



Q-Q PLOT FOR MODEL I:

- The points doesn't seem to fall directly in a line.
- Distribution is highly abnormal, indicating this is a bimodal distribution
- Sample data show skewness to the right.
- Possible reasoning is due to the extreme values in the dataset

04.
CONCLUSION



CONCLUSION OF STUDY

- Conclusions:
  - In our association analysis, we found that social media usage is tied to thoughts of feeling lonely & depressed.
  - In our regression analysis, we found that the third model is explaining that if we try to explain sleeping quality only with the variable distraction level by user and feelings of comparison to others there is a significant relationship between them - the level of distraction for using social media and the emotions related to it can explain in some relation the predicted variable.