Final Paper - Part 1
Group 7

- *your specific research questions*

  Does work-life balance play an important role in the attrition rate of employees? If yes, how impactful is it on their decision? Are there other factors that contribute to the attrition rate of employees?
  Dependent variable: Attrition Rate
  Independent variable: Work Life Balance, Environment Satisfaction, Education Level, *Gender, and Marital Status*

- *brief description of your dataset. Be sure to mention the unit of analysis in your dataset, as well as the number of observations in the dataset (the "n" about whom there is information for your key variables)*

  Unit of analysis: Individual employees
  No. of observations: 1676
  Type of variables: All our independent variables are categorical, and our dependent variable is Boolean (Yes/No).
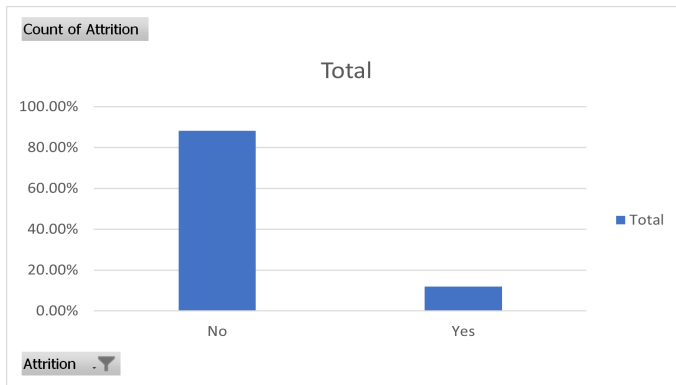
- *description of your dependent variable and descriptive statistics for it*

  Attrition is one main factor that can demolish even the most successful and stable organizations in a startlingly short period of time. While many businesses focus on the costly process of hiring and recruiting, there isn't enough focus on addressing the issues that cause top talent to leave. Understanding the motivations behind employee attrition is the critical first step in understanding and properly addressing the problem.

  Below is the overview of each of the attrition rate with its descriptive statistics that is laying out the frequency, % of total along with the histogram to display the same.

Row Labels	% of Attrition
No	88.13%
Yes	11.87%

Row Labels	Count of Attrition
No	1477
Yes	199

- *description of your independent variable(s) and descriptive statistics for them*

We are trying to compare and reveal the true drivers of attrition, by taking into account primary factors of our analysis such as Work life Balance and other factors such as Environment Satisfaction, Education Level, *Gender and Marital Status.* Below is the overview of each of the independent variables with the help of their descriptive statistics that is laying out their frequency, % of total with respect to the dependent variable along with the histogram to display the same.
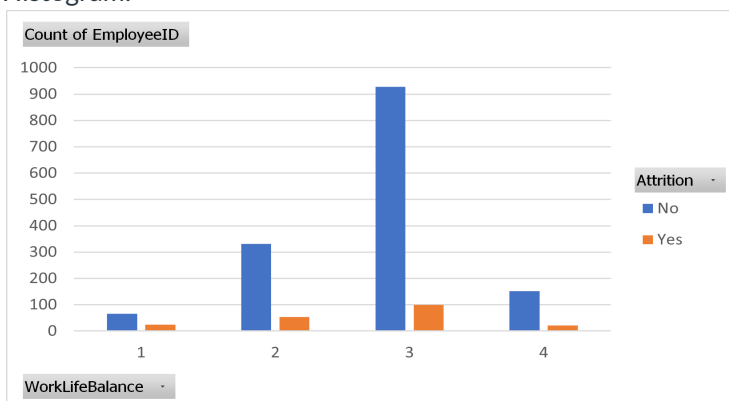
Descriptive Statistics:

IV 1: Work life Balance

Frequency for Work life Balance by Attrition rate:

Row Labels	1	2	3	4
No	66	331	928	152
Yes	24	54	100	21

% for Work life Balance by Attrition rate:

Row Labels	1	2	3	4
No	3.94%	19.75%	55.37%	9.07%
Yes	1.43%	3.22%	5.97%	1.25%

Histogram:

IV 2: Environment Satisfaction

Frequency for Environment Satisfaction by Attrition rate:

Row Labels	1	2	3	4
No	266	288	464	459
Yes	64	38	48	49

% for Environment Satisfaction by Attrition rate:

Row Labels	1	2	3	4
No	15.87%	17.18%	27.68%	27.39%
Yes	3.82%	2.27%	2.86%	2.92%

Histogram:
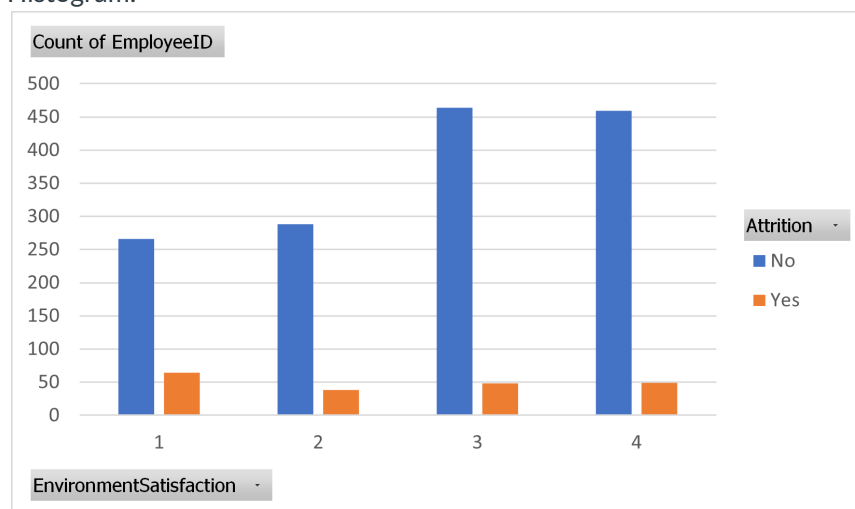
IV 3: Education

Frequency for Education by Attrition rate:

Row Labels	1	2	3	4	5
No	169	287	566	400	55
Yes	27	35	89	47	1

% for Education by Attrition rate:

Row Labels	1	2	3	4	5
No	10.08%	17.12%	33.77%	23.87%	3.28%
Yes	1.61%	2.09%	5.31%	2.80%	0.06%

Histogram –
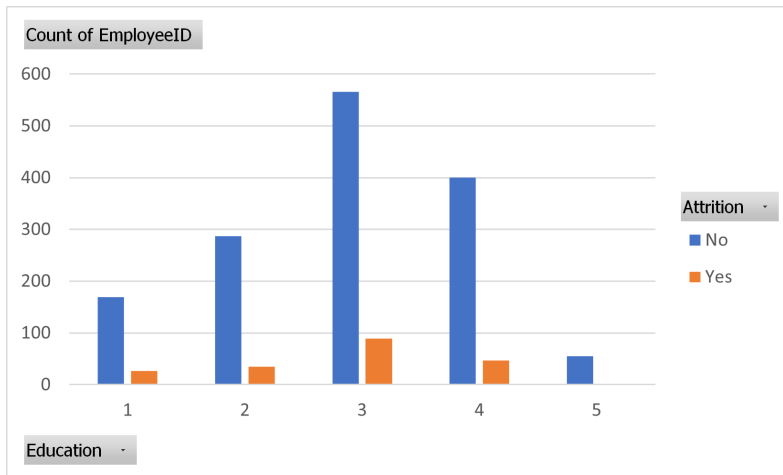
IV 4: Gender

Frequency for Gender by Attrition rate:

Row Labels	Female	Male
No	592	885
Yes	86	113

% for Gender by Attrition rate:

Row Labels	Female	Male
No	35.32%	52.80%
Yes	5.13%	6.74%

Histogram:
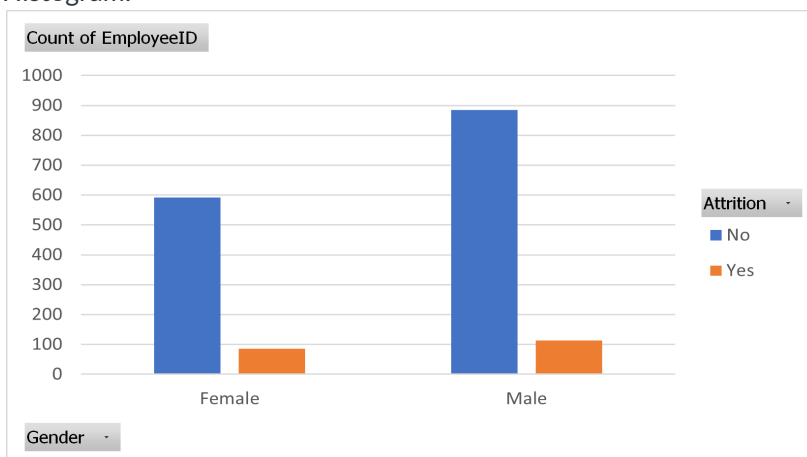
IV 5: Marital Status

Frequency for Marital Status by Attrition rate:

Row Labels	Divorced	Married	Single

No	353	716	408
Yes	24	61	114

% for Marital Status by Attrition rate:

Row Labels	Divorced	Married	Single
No	21.06%	42.72%	24.34%
Yes	1.43%	3.64%	6.80%

Histogram:
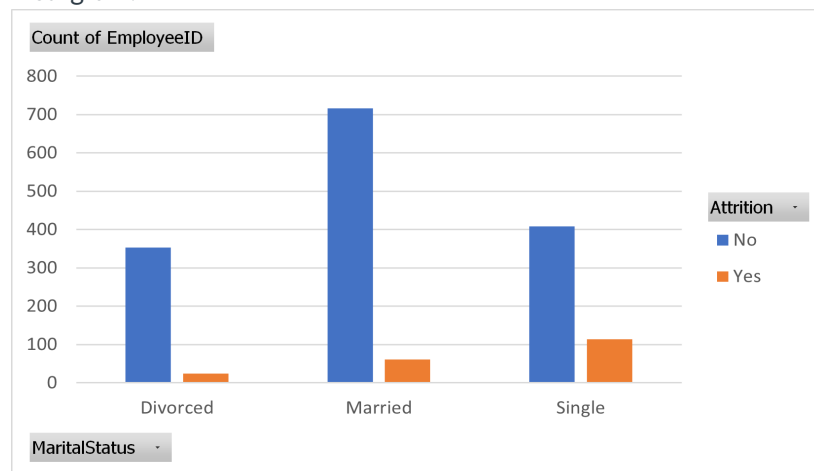
- *preliminary/rough answer to your research question*

  We use the chi square test to compare the independent variables iwork-life balance, environmental satisfaction, gender, marital status, and education with the dependent variable attrition rate. We try to find answers to a few questions –
  a. whether employees with poor work-life balance are more likely to quit.
  b. whether Employees who leave their jobs are more likely to be women than men.
  c. whether women who are married have a higher attrition compared to women who are unmarried
  d. whether employees with a lower level of education leave the company more.
  e. whether employees with a higher environmental satisfaction have a lower attrition

  After conducting the chi square test between work life balance and attrition, we found that the p value (X^2=25, Df = 3) is 1.54405e-05. Since p-value is less than 0.05, we rejected the null hypothesis which stated that there is no relationship between work-life balance and attrition rate.

  We hope to conclude that there will be a relationship between all our other independent variables and our dependent variable, attrition rate as well. Further we would want to analyze the impact of each of the IV-DV relationships using logistic regression.