**Executive Summary**

**DDSAnalytics Case Study #2**

**DataBusters Project Team: Griffin Lane, Josie McDaniel, Kris Henley**

**Introduction**

For this study DataBusters has been asked to conduct exploratory data analysis (EDA) to help DDSAnalytics determine leading factors that can lead to higher attrition. The project team will also explore existing trends and observations based on the data provided that may prove useful in evaluating attrition.

**Data Description**

\*The data set used for the purposes of this analysis was provided by DDSAnalytics CEO Dr. Faizen Javed.

\*The data consists of 1470 observations containing individual employment details that cover a wide range of explanatory variables (35 in total).

\*These variables differ in class and are mixed by factor, character, and integer respectively.

\*For the purposes of this analysis we will treat the Attrition variable as our dependent variable.

**Results**

Through extensive exploratory analysis (EDA) using advanced linear regression techniques, and scientifically supported measures, the DataBusters project team was able to conclusively identify the top factors that we believe are currently contributing to increased attrition rates based on the data provided. Our team was also able to identify factors that we believe could help DDSAnalytics identify situations where attrition is less likely to occur. Here are the results:

**Strong Associations with Attrition:**

1. **Distance From Home** – The average commuting distance for the attrition group was 9.2 miles. Although this is comparable to the non-attrition group, our data suggests that when paired with one or more of the other factors below commute time can contribute to increased attrition frequencies.
2. **Frequent Travel** - About 30% of our attrition group travelled frequently for business, as compared to only 17% in our non-attrition group.
3. **Job Changers** – This one is fairly obvious. Individuals with a higher number of total companies work for are more likely to be lost to attrition.
4. **Overtime** - 54% of our attrition group reported having to work overtime, as compared to 9% with our non-attrition group.

**Strong Associations against Attrition:**

1. In general DataBusters shows that employees with educations relating to the **Medical or Life Sciences fields** tend to have lower attrition rates.
2. Our non-attrition group showed a 2.7 average **Environment Satisfaction** score. While this again is comparable to our attrition group’s score, our data suggests that when this score is factored into a comparison with one or more of the other variables in this list attrition will be lower.
3. Our data also suggests that higher numbers in **Job Involvement** scores and **Years in Current Role** suggest lower attrition rates.

**Other Trends and Observations**

1. Job roles that seem to show a very high attrition rate when compared against the others in this data set are Laboratory Technician and Sales Representative.
2. Since most Laboratory Technicians will likely be formally educated in the Life Sciences the fact that they show a higher likelihood of attrition is a bit unique in the Medical or Life Science fields. This particular role should be examined more closely with particular attention paid to the other contributing factors to attrition.

\*\*BONUS\*\*

The DataBusters Project Team utilized the regression model we formulated to provide the analysis above, and used it to produce a preliminary predictive model. When paired against our test data set we are happy to report we were able to see an **88% success rate in predicting attrition!**

\*Please note that the results above pertain only to the test sample provided. Additional controlled data collection and analysis would need to be conducted in a more formal and randomized experiment in order to establish definitive causal relationships between the variables in question and the outcome recorded.