

Normality Test

Data Analyst: Gyro A. Madrona Department: Electrical Engineering

```
1  # imports and packages
2  import pandas as pd
3  import matplotlib.pyplot as plt
4  from scipy import stats
```

Shapiro-Wilk Test

Data Cleaning

```
1  # check for outliers
2  plt.boxplot(df['Defects'],tick_labels=['Defects'])
3  plt.show()

1  # shapiro-wilk test
2  w_stat, p_value = stats.shapiro(df['Defects'])

Python
```

Aderson-Darling Test

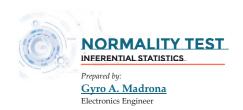
```
1 # load dataset
2 df = pd.read_csv(r"raw\defects-count.csv",delimiter=",")
3 df

[] 娜Open'df in Data Wrangler

1 # summary of dataframe
2 df.info()

1 # summary of statistics
2 df.describe()

Python
```



Data Cleaning

```
1  # check for outliers
2  plt.boxplot(df['Defects'],tick_labels=['Defects'])
3  plt.show()

1  # anderson-darling statistics
2  a2_stat, critical, alpha = stats.anderson(df['Defects'])

Python
```