ANOVA

INFERENTIAL STATISTICS

Prepared by:

Gyro A. Madrona

Electronics Engineer

ANOVA

Data Analyst: Gyro A. Madrona
Department: Electrical Engineering

```python
1  #%pip install pingouin --upgrade
```

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

```python
1  # load dataset
2  df = pd.read_csv(r"raw\hydroelectric-sample.csv",
3                   delimiter=',')
4  df
```
Open 'df' in Data Wrangler

```python
1  # summary of dataframe
2  df.info()
```

```python
1  # summary of statistics
2  df.describe()
```

Data Cleaning

```python
1  # check for 3-sigma outliers
2  plt.boxplot(df[['Plant_1','Plant_2','Plant_3']],
3              tick_labels=df.columns)
4  plt.title("Hyrdroelectric Plants")
5  plt.ylabel("Production MWh")
6  plt.show()
```
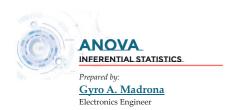
Normality Test

```python
1  # load dataset
2  df = pd.read_csv(r"raw\hydroelectric-sample.csv")
3  df
```
Open 'df' in Data Wrangler

```python
1  # summary of dataframe
2  df.info()
```

```python
1  # summary of statistics
2  df.describe()
```

ANOVA
INFERENTIAL STATISTICS

Prepared by:
Gyro A. Madrona
Electronics Engineer

One-Way ANOVA

```python
1  # hypothesis
2  # Ho:equal means
3  # Ha: at least one is not equal
4
5  f_stat, p_value = stats.f_oneway(df['Plant_1'],
6                                    df['Plant_2'],
7                                    df['Plant_3'])
8  p_value
```

Python