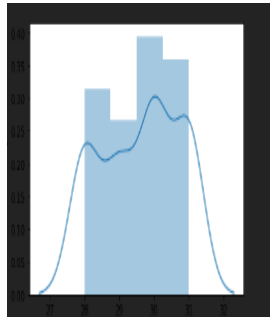
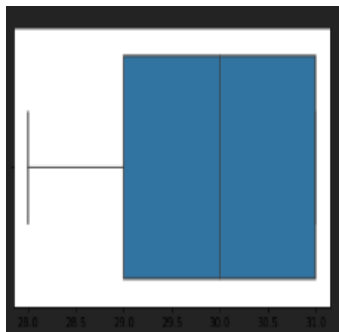


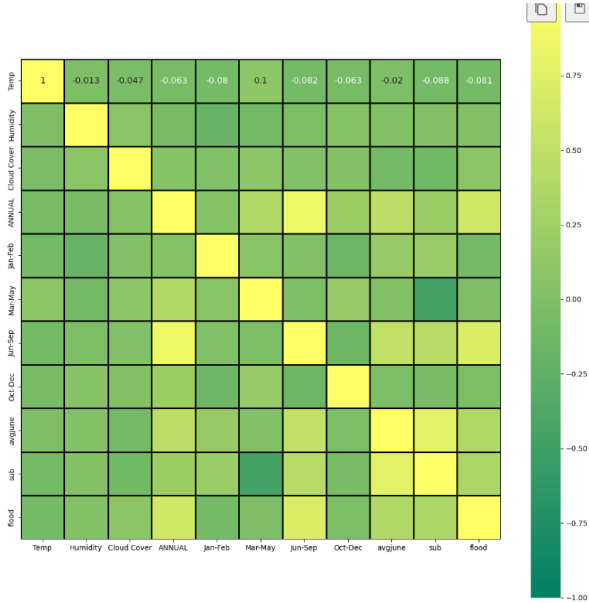
Data Collection and Preprocessing Phase

| | |
|---------------|--|
| Date | 10 July 2024 |
| Team ID | 739726 |
| Project Title | Rising Waters: A Machine Learning Approach To Flood Prediction |
| Maximum Marks | 6 Marks |

Data Exploration and Preprocessing Template

Identifies data sources, assesses quality issues like missing values and duplicates, and implements resolution plans to ensure accurate and reliable analysis.

| Section | Description | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|----------------------|---|------|----------|-------------|-------------|---------|---------|---------|------------|------------|---------|---|----|----|------|--------|------|-------|--------|-------|------------|---|----|----|------|--------|-----|-------|--------|-------|------------|---|----|----|------|--------|------|-------|--------|-------|------------|---|----|----|------|--------|------|-------|--------|-------|------------|---|----|----|------|--------|------|-------|--------|-------|------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|----|------|--------|------|-------|--------|-------|------------|-----|----|----|------|--------|------|-------|--------|-------|------------|-----|----|----|------|--------|------|-------|--------|-------|------------|-----|----|----|------|--------|------|-------|--------|-------|------------|-----|----|----|------|--------|-----|-------|--------|-------|------------|
| Data Overview | Dimensions 115rows x 9columns | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | <table><tr><th></th><th>Temp</th><th>Humidity</th><th>Cloud Cover</th><th>ANNUAL</th><th>Jan-Feb</th><th>Mar-May</th><th>Jun-Sep</th><th>Oct-Dec</th><th>avgjune</th></tr><tr><td>0</td><td>29</td><td>70</td><td>30.0</td><td>3248.6</td><td>73.4</td><td>386.2</td><td>2122.8</td><td>666.1</td><td>274.866667</td></tr><tr><td>1</td><td>28</td><td>75</td><td>40.0</td><td>3326.6</td><td>9.3</td><td>275.7</td><td>2403.4</td><td>638.2</td><td>130.300000</td></tr><tr><td>2</td><td>28</td><td>75</td><td>42.0</td><td>3271.2</td><td>21.7</td><td>336.3</td><td>2343.0</td><td>570.1</td><td>186.200000</td></tr><tr><td>3</td><td>29</td><td>71</td><td>44.0</td><td>3129.7</td><td>26.7</td><td>339.4</td><td>2398.2</td><td>365.3</td><td>366.066667</td></tr><tr><td>4</td><td>31</td><td>74</td><td>40.0</td><td>2741.6</td><td>23.4</td><td>378.5</td><td>1881.5</td><td>458.1</td><td>283.400000</td></tr><tr><td>...</td><td>...</td><td>...</td><td>...</td><td>...</td><td>...</td><td>...</td><td>...</td><td>...</td><td>...</td></tr><tr><td>110</td><td>28</td><td>71</td><td>30.0</td><td>3035.1</td><td>66.2</td><td>313.5</td><td>2209.1</td><td>446.3</td><td>262.833333</td></tr><tr><td>111</td><td>29</td><td>71</td><td>37.0</td><td>2151.1</td><td>18.3</td><td>287.4</td><td>1535.6</td><td>309.8</td><td>143.433333</td></tr><tr><td>112</td><td>30</td><td>74</td><td>42.0</td><td>3255.4</td><td>43.9</td><td>218.5</td><td>2561.2</td><td>431.8</td><td>347.566667</td></tr><tr><td>113</td><td>31</td><td>71</td><td>31.0</td><td>3046.4</td><td>14.9</td><td>364.5</td><td>2164.8</td><td>502.1</td><td>151.466667</td></tr><tr><td>114</td><td>28</td><td>71</td><td>34.0</td><td>2600.6</td><td>8.9</td><td>465.9</td><td>1514.7</td><td>611.1</td><td>187.866667</td></tr></table> | | Temp | Humidity | Cloud Cover | ANNUAL | Jan-Feb | Mar-May | Jun-Sep | Oct-Dec | avgjune | 0 | 29 | 70 | 30.0 | 3248.6 | 73.4 | 386.2 | 2122.8 | 666.1 | 274.866667 | 1 | 28 | 75 | 40.0 | 3326.6 | 9.3 | 275.7 | 2403.4 | 638.2 | 130.300000 | 2 | 28 | 75 | 42.0 | 3271.2 | 21.7 | 336.3 | 2343.0 | 570.1 | 186.200000 | 3 | 29 | 71 | 44.0 | 3129.7 | 26.7 | 339.4 | 2398.2 | 365.3 | 366.066667 | 4 | 31 | 74 | 40.0 | 2741.6 | 23.4 | 378.5 | 1881.5 | 458.1 | 283.400000 | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | 110 | 28 | 71 | 30.0 | 3035.1 | 66.2 | 313.5 | 2209.1 | 446.3 | 262.833333 | 111 | 29 | 71 | 37.0 | 2151.1 | 18.3 | 287.4 | 1535.6 | 309.8 | 143.433333 | 112 | 30 | 74 | 42.0 | 3255.4 | 43.9 | 218.5 | 2561.2 | 431.8 | 347.566667 | 113 | 31 | 71 | 31.0 | 3046.4 | 14.9 | 364.5 | 2164.8 | 502.1 | 151.466667 | 114 | 28 | 71 | 34.0 | 2600.6 | 8.9 | 465.9 | 1514.7 | 611.1 | 187.866667 |
| | | Temp | Humidity | Cloud Cover | ANNUAL | Jan-Feb | Mar-May | Jun-Sep | Oct-Dec | avgjune | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | 0 | 29 | 70 | 30.0 | 3248.6 | 73.4 | 386.2 | 2122.8 | 666.1 | 274.866667 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | 1 | 28 | 75 | 40.0 | 3326.6 | 9.3 | 275.7 | 2403.4 | 638.2 | 130.300000 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | 2 | 28 | 75 | 42.0 | 3271.2 | 21.7 | 336.3 | 2343.0 | 570.1 | 186.200000 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | 3 | 29 | 71 | 44.0 | 3129.7 | 26.7 | 339.4 | 2398.2 | 365.3 | 366.066667 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | 4 | 31 | 74 | 40.0 | 2741.6 | 23.4 | 378.5 | 1881.5 | 458.1 | 283.400000 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| | 110 | 28 | 71 | 30.0 | 3035.1 | 66.2 | 313.5 | 2209.1 | 446.3 | 262.833333 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 111 | 29 | 71 | 37.0 | 2151.1 | 18.3 | 287.4 | 1535.6 | 309.8 | 143.433333 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 112 | 30 | 74 | 42.0 | 3255.4 | 43.9 | 218.5 | 2561.2 | 431.8 | 347.566667 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 113 | 31 | 71 | 31.0 | 3046.4 | 14.9 | 364.5 | 2164.8 | 502.1 | 151.466667 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 114 | 28 | 71 | 34.0 | 2600.6 | 8.9 | 465.9 | 1514.7 | 611.1 | 187.866667 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| 115 rows x 9 columns | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| Univariate Analysis |  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
| |  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

| | |
|-------------------------------------|---|
| Bivariate Analysis | - |
| Multivariate Analysis |  |
| Outliers and Anomalies | - |
| Data Preprocessing Code Screenshots | |
| Loading Data | <pre>data=pd.read_excel("flood dataset.xlsx")</pre> |
| Handling Missing Values | <pre>data.isnull().sum() Temp False Humidity False Cloud Cover False ANNUAL False Jan-Feb False Mar-May False Jun-Sep False Oct-Dec False avgJune False sub False flood False dtype: bool</pre> |
| Feature Scaling | <pre>#import StandardScaler from sklearn.preprocessing import StandardScaler #create object to StandardScaler class sc=StandardScaler() x_train=sc.fit_transform(x_train) x_test=sc.fit_transform(x_test)</pre> <pre>#import dump class from joblib from joblib import dump dump(sc,"transform.save")</pre> |

| | |
|---------------------|------------------------------------|
| Feature Engineering | Attached code in final submission. |
| Save Processed Data | - |