## Scala's real-world project repository data

**Functies:**

1. import pandas as pd <https://www.statology.org/import-pandas-as-pd/>
2. pd.read\_csv <https://pythoncursus.nl/open-csv-python/>

## Preparing and cleaning the data

**Functies:**

1. pd.concat <https://pandas.pydata.org/docs/reference/api/pandas.concat.html>
2. Pandas.to\_datetime() <https://www.geeksforgeeks.org/python-pandas-to_datetime/>

## Merging the DataFrames

**Functies:**

1. .merge <https://pandas.pydata.org/docs/reference/api/pandas.merge.html>

## Is the project still actively maintained?

**Functies:**

1. %matplotlib inline <https://www.scaler.com/topics/matplotlib/matplotlib-inline/>
2. dt.month / dt.year <https://towardsdatascience.com/dealing-with-dates-in-pythons-dataframe-part-2-the-basics-9ad5edacd2f8>
3. groupby <https://realpython.com/pandas-groupby/>
4. count <https://www.simplilearn.com/tutorials/python-tutorial/count-in-python#:~:text=Count()%20is%20a%20Python,list%2C%20as%20the%20name%20implies>.

## Is there camaraderie in the project?

**Functies:**

1. %matplotlib inline <https://www.scaler.com/topics/matplotlib/matplotlib-inline/>
2. groupby <https://realpython.com/pandas-groupby/>
3. count <https://www.simplilearn.com/tutorials/python-tutorial/count-in-python#:~:text=Count()%20is%20a%20Python,list%2C%20as%20the%20name%20implies>
4. hist() <https://matplotlib.org/stable/api/_as_gen/matplotlib.pyplot.hist.html>
5. bins= <https://stackoverflow.com/questions/51491391/meaning-of-bins-in-python>

## What files were changed in the last ten pull requests?

**Functies:**

1. nlargest() <https://www.w3schools.com/python/pandas/ref_df_nlargest.asp#:~:text=Definition%20and%20Usage,value%20for%20a%20specified%20column>.
2. merge <https://pandas.pydata.org/docs/reference/api/pandas.merge.html>
3. Set() <https://www.w3schools.com/python/python_sets.asp>
4. Print() <https://www.w3schools.com/python/ref_func_print.asp#:~:text=The%20print()%20function%20prints,before%20written%20to%20the%20screen>

## Who made the most pull requests to a given file?

**Functies:**

1. groupby <https://realpython.com/pandas-groupby/>
2. count <https://www.simplilearn.com/tutorials/python-tutorial/count-in-python#:~:text=Count()%20is%20a%20Python,list%2C%20as%20the%20name%20implies>
3. sort\_values <https://www.w3schools.com/python/pandas/ref_df_sort_values.asp>
4. head()
5. Print

<https://www.w3schools.com/python/ref_func_print.asp#:~:text=The%20print()%20function%20prints,before%20written%20to%20the%20screen>

## Who made the last ten pull requests on a given file?

**Functies:**

1. merge <https://pandas.pydata.org/docs/reference/api/pandas.merge.html>
2. nlargest() <https://www.w3schools.com/python/pandas/ref_df_nlargest.asp#:~:text=Definition%20and%20Usage,value%20for%20a%20specified%20column>
3. Set() <https://www.w3schools.com/python/python_sets.asp>
4. Print()

<https://www.w3schools.com/python/ref_func_print.asp#:~:text=The%20print()%20function%20prints,before%20written%20to%20the%20screen>

## The pull requests of two special developers

**Functies:**

1. %matplotlib inline <https://www.scaler.com/topics/matplotlib/matplotlib-inline/>
2. isin() <https://www.w3schools.com/python/pandas/ref_df_isin.asp#:~:text=Definition%20and%20Usage,the%20specified%20values%2C%20otherwise%20False%20>
3. groupby <https://realpython.com/pandas-groupby/>
4. agg() <https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.agg.html>
5. reset\_index <https://www.w3schools.com/python/pandas/ref_df_reset_index.asp>
6. pivot\_table <https://pandas.pydata.org/docs/reference/api/pandas.pivot_table.html>
7. plot() <https://www.activestate.com/resources/quick-reads/how-to-display-a-plot-in-python/>

## Visualizing the contributions of each developer

**Functies:**

1. isin() <https://www.w3schools.com/python/pandas/ref_df_isin.asp#:~:text=Definition%20and%20Usage,the%20specified%20values%2C%20otherwise%20False%20>
2. groupby <https://realpython.com/pandas-groupby/>
3. count <https://www.simplilearn.com/tutorials/python-tutorial/count-in-python#:~:text=Count()%20is%20a%20Python,list%2C%20as%20the%20name%20implies>
4. reset\_index <https://www.w3schools.com/python/pandas/ref_df_reset_index.asp>
5. pivot\_table <https://pandas.pydata.org/docs/reference/api/pandas.pivot_table.html>
6. plot() <https://www.activestate.com/resources/quick-reads/how-to-display-a-plot-in-python/>