User Manual

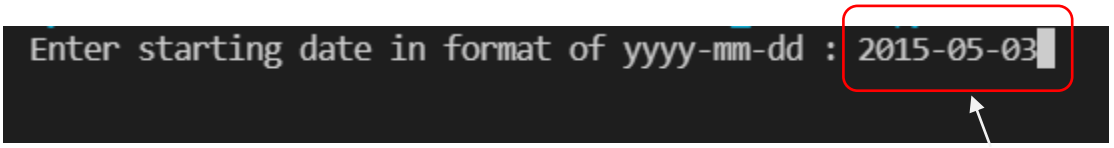
Initial Requirements

- You need python 3 in your computer to run these .py files. You can download python here.
- You also need to install pandas library. You can use these two [01 or 02] websites to install pandas.
- Then you need matplotlib library for visualization. Here you can find way to install it.

Feature 01

(For a user-selected period, display the information of all accidents that happened in the period.)

- First of all, you have to run this ***first_feature.py*** file. There will be four inputs. You have to read the banner of the input and type in the correct format.
    - 1st one is the starting date for the period that you wish.
    - You have to type with correct format of date otherwise error will occur.
    - Correct format is
        - 4-digit year (2015)–2-digit month (05)–2-digit day (03)
        - No spaces are allowed between digits and ' - '(Hyphen Mark)
        - Then press enter



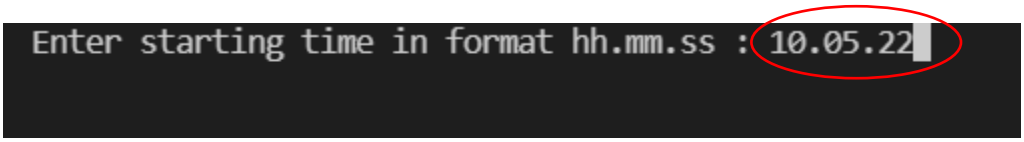
1st input prompt with example input

    - Then you have to enter the ending date for your period. Procedure as same as above.

- In here please your ending date does not including the output results. It just output the results up to ending date. (That means your code shows the output excluding the ending date.) If you want to fix this (to include the ending date too.) you have to change the operator in the python code as follows. (Code line no 27)

```
# filter out the selected range of dates.....................................
'''
this takes upto end date. does not take end date.
If you want to consider end date too.. simply change the '<' symbol with '<='
'''
filt = (df['DATE'] >= pd.to_datetime(s_date)) & (df['DATE'] < pd.to_datetime(e_date))
```

- Then you have to enter the starting hour in 24-hour format for selection as the third input after entering the correct ending date.
- Correct format is
  - 2-digit hour (10).2-digit minute (05).2-digit second (22)
  - No spaces are allowed between digits and ' . '(Full Stop Mark)
  - Then press the enter.



- Follow the same procedure for the ending time. You have to consider about excluding and including effect as above. And you can simply edit the code accordingly as mentioned above. (Code line no 32)

- Your results according to the selection will save in the html file at the location that you provided. You have to change the location for that file as you wish. (Code Line no 42)

```
# view the resultant data table from this html page.....................
new_df.to_html('C:\\Users\\Documents\\accident_results_over_period.html')
```

- Use correct directory for the file as mentioned in red box (*Use double backslashes*)
- Enter the file name with .html extension as mentioned in yellow box.
- Then go to the location and open the html file. You can see the resultant outputs for selected period.

Special notes

- *We use few selected columns among all columns to show the details. If you want to change the columns as you wish, simple change the column no that you want in the code line no 07.*

```
# read the input .csv file for processing purposes..you can choose columns as you wish..
df = pd.read_csv("CrashStatisticsVictoria.csv", usecols=[4,5,6,7,8,11,15,26])
```

- *Then our algorithm is used 00.00.00 hr as the start of the day (Just after the midnight) and 24.00.00 hr as the end of the day(midnight).*

Feature 02

(For a user-selected period, produce a chart to show the number of accidents in each hour of the day (on average))

- First of all, you have to run this ***second_feature.py*** file. There will be four inputs. You have to read the banner of the input and type in the correct format. Use the instruction of Feature 01 for that.
- You can change the excluding/including date or time as feature 01 method.
- Then you will be able to see the average accidents over hour for given day period as bar graph.
- If you want to see the results for whole day you have to enter 00.00.00 as the starting time and 24.00.00 as the ending time.
- Otherwise, resultant graph only shows the selected hour period.

Feature 03

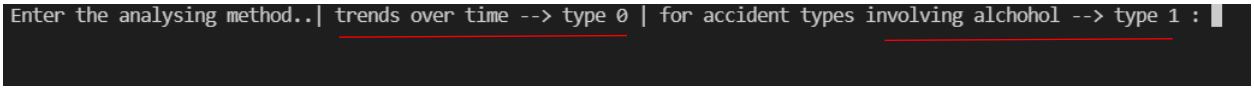
(For a user-selected period, retrieve all accidents caused by an accident type that contains a keyword (user entered), e.g., collision, pedestrian.)

- First of all, you have to run this ***third_feature.py*** file. There will be four inputs. You have to read the banner of the input and type in the correct format. Use the instruction of Feature 01 for that.
- You can change the excluding/including date or time as feature 01 method.
- As an extra input, user have to enter the keyword for searching.
- Make sure to enter the correctly formatted keyword as the data set shows under the "ACCIDENT_TYPE" column. (Use uppercase for 1st letter)
- No need to enter the entire word. Just few letters with correct order is enough.
- Example Keywords are,
    - Ped - Struck Pedestrian
    - Stru - Struck Pedestrian
    - Stru - Struck animal
    - Pedestrian - Struck Pedestrian
    - Coll - Collision with vehicle
    - Coll - Collision with a fixed object
    - Coll – No output data will be showed (since there are no such keyword, incorrect format)
- If you enter a few letters with correct format, it will output the entire results that have input letters. Just filtered the entered word.
- *If you want to filter out exact details under one keyword you have to enter the full word as keyword.*
- Then the results will save as html file at given location. You can follow the same instruction under the feature 01 for save the file and open the html file.

Feature 04

(Allow the user to analyze the impact of alcohol in accidents – i.e.: trends over time, accident types involving alcohol.)

- First of all, you have to run this ***fourth_feature.py*** file. There is no period selection as others. It is only asked for an integer input (0 or 1) as the category that you wish to filter the results under the alcoholic effect.
- If you wish to see the results of alcoholic effect over the time period, you can simply enter "0" as the input or for the results of accident types involving alcohol, simply enter the "1" as input.
- For the other inputs, you will see the message. It tells to refer the user manual.
- When you run the file prompt also ask for the correct input with a banner.



- Then you will be able to see the resultant outputs under your selection as a bar charts...

Feature 05

(One other 'insight' or analysis tool of your choice)

- First of all, you have to run this ***fifth_feature.py*** file. There will be four inputs. You have to read the banner of the input and type in the correct format. Use the instruction of Feature 01 for that.
- You can change the excluding/including date or time as feature 01 method.
- As the results of this python script, you will be able to see the number of accidents for given period w.r.t the speeds.