**FRS Dataset**

**Introduction:**

Using flight fare price prediction (AI tool) we are going to predict different flight prices with best services and cheapest price.

First I am going to extract a real time fight data, including flowing columns: (Not going to use Kaggle.com)

* Flight name
* Flight path
* Origin
* Origin Local Code
* Destination
* Destination Local Code
* Departure Date
* Return Date
* Flight Type
* Countries

I am extracting the flight real time data using automation, technique to automate tasks, using python.

Going to use selenium for data extraction from website name hotwire.com website.

**Main Points:**

* All flights from Pakistan to other countries are extracted. **(Islamabad – London, Peshawar – UAE etc.).**
* We have multiple airports in Pakistan, so extracting data for each airport.
* There are 100+ Airports in the world, but because of data quantity and execution time, I am only considering 13 most popular countries (Airports).
* Data is extracted within Date **05/10/2023** to **05/20/2023** (10 May to 20 May). Because of large number of data and execution time. **(3000 entries took about 5 hours).**
* Screenshots were added at bottom for better understanding.

Here are the steps I have done during data extraction:

**Step1:**  Preparing pre work data (Files)

**Step2:** Writing Python Script using Selenium (Automation Tool).

**Step3:** Attaching my pre-build module to extract data anonymously (using proxies and user agents). Pre-build module is on my [GitHub Account](https://github.com/HammadRafique29). (Verification)

**Step4:** Saving extracted data in csv and text file for usage.

**Screenshot’s:**













