### Using the call detail record layout (Cdr.xlsx) to create dataset containing 1000 rows/call. The data is two have no more than 30 customers i.e. uid. The customers are located in Tunisia. Use the total\_time column and the “Analysis of Abnormal Call Duration” and Example statement to make some records suspect by placing a “1” in the fraud column. No more than 10% of cause are to be flagged as fraud.

### Analysis of Abnormal Call Duration:

Fraud detection may involve analyzing extremely short or abnormally long call durations, which may indicate fraudulent activity such as SIM card fraud.

### Example : A large number of very short calls or several prolonged calls to the same number may indicate a fraud attempt.

**ChatGPT Prompts Session**

1. Here is the requirement. Using the call detail record layout (Cdr.xlsx) to create dataset containing 100,000 rows/call. The data is to have no more than 30 customers i.e. caller phone. And 30 destination phone The customers are located in Tunisia. Both phone numbers should be like the phone numbers in Tunisia.

The call dates are between 2023-10-01 and 2023-12-31. The call dates are to be random but there should more calls on Saturday and Sunday than during the rest of day the week.

1. The call\_duration is a random series of numbers between 0.25 and 180 minutes.
2. A fraud call is any call that is less than 2 minutes and 160 minutes ”. There is a column called fraud which will contain a “F” if the call is a fraud, and “G” if the call is good.
3. No more than 10% of calls are to be flagged as fraud.
4. Finaly 0.05 percent are to be “false positives”, and other 0.05 percent are to be false “negative”,
5. Ok, write data to a csv file..