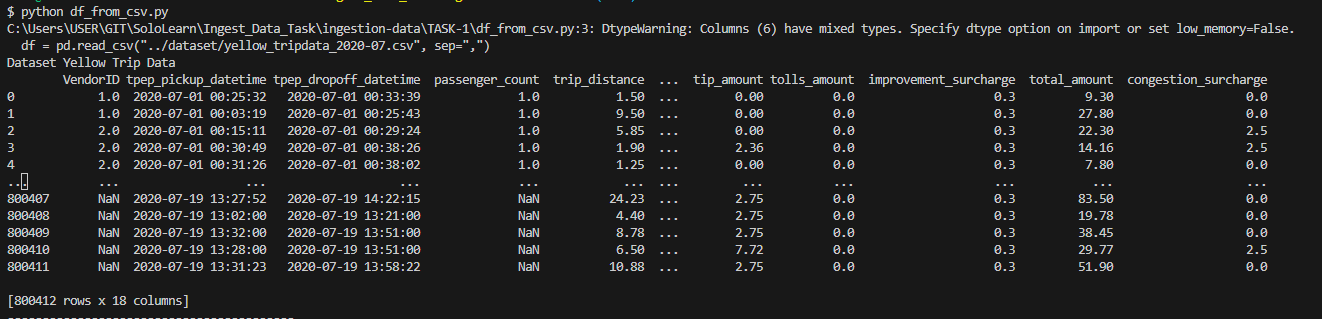
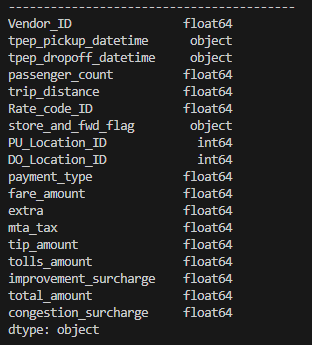
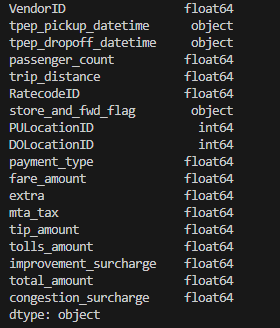
Nurul Taufik Kurrahman

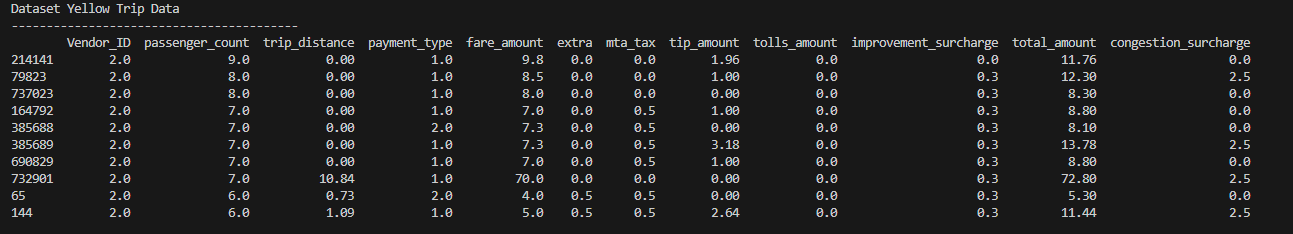
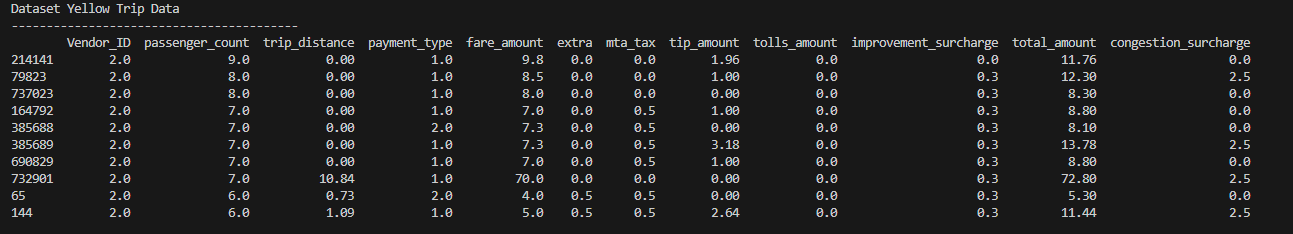
**Task 1**

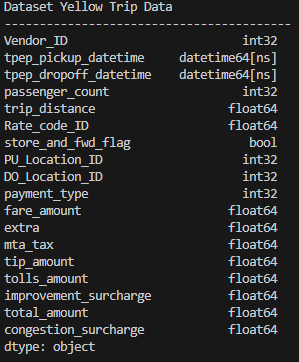
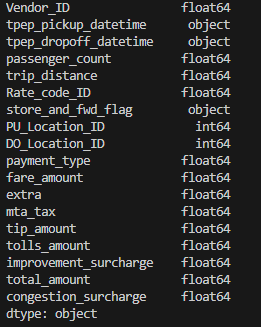
1. We have already learned how to create DataFrame from files here. Now, we are going to create a DataFrame from a larger csv file on our datasets.

[](https://github.com/ALTA-DE1-CHARIS-14101996/ingestion-data/blob/main/TASK-1/image-1.png)

1. Rename all the columns with snake\_case format.

[](https://github.com/ALTA-DE1-CHARIS-14101996/ingestion-data/blob/main/TASK-1/image-5.png)[](https://github.com/ALTA-DE1-CHARIS-14101996/ingestion-data/blob/main/TASK-1/image-3.png)\*Before Rename. \*After Rename.

1. [](https://github.com/ALTA-DE1-CHARIS-14101996/ingestion-data/blob/main/TASK-1/image-7.png)Select only 10 top of highest number of passenger\_count, show only columns vendor\_id, passenger\_count, trip\_distance, payment\_type, fare\_amount, extra, mta\_tax, tip\_amount, tolls\_amount,improvement\_surcharge, total\_amount, congestion\_surcharge from the DataFrame.
2. [Extra] Cast the data type to the appropriate value.

[](https://github.com/ALTA-DE1-CHARIS-14101996/ingestion-data/blob/main/TASK-1/image-10.png)[](https://github.com/ALTA-DE1-CHARIS-14101996/ingestion-data/blob/main/TASK-1/image-8.png)\*Before Change Data Type \*After Change Data Type