

# Data Intake Report

Name: G2M insight for Cab Investment firm

Report date: 9/15/2023

Internship Batch: LISUM25

Version: 1.0

Data intake by: Haseeb Javed

Data intake reviewer: Haseeb Javed

Data storage location: <https://github.com/haseebjaved4652/DataSet>

## Tabular data details:

### Hypothesis #1

<b>Total number of observations</b>	2
<b>Total number of files</b>	1
<b>Total number of features</b>	
<b>Base format of the file</b>	ipynb
<b>Size of the data</b>	KB

### Hypothesis #2

<b>Total number of observations</b>	3
<b>Total number of files</b>	1
<b>Total number of features</b>	
<b>Base format of the file</b>	ipynb
<b>Size of the data</b>	60 KB

### Hypothesis #3

<b>Total number of observations</b>	2
<b>Total number of files</b>	3
<b>Total number of features</b>	
<b>Base format of the file</b>	ipynb
<b>Size of the data</b>	53 KB

### Hypothesis #4

<b>Total number of observations</b>	2
<b>Total number of files</b>	1
<b>Total number of features</b>	
<b>Base format of the file</b>	ipynb
<b>Size of the data</b>	86 KB

#### Hypothesis #5

<b>Total number of observations</b>	3
<b>Total number of files</b>	1
<b>Total number of features</b>	
<b>Base format of the file</b>	ipynb
<b>Size of the data</b>	116 KB

**Note: Replicate same table with file name if you have more than one file.**

**Proposed Approach:**

- Mention approach of dedup validation (identification)
- Mention your assumptions (if you assume any other thing for data quality analysis)

**Note: Convert this doc in pdf and provide the link of pdf file in your dashboard.  
Please do not forget to remove this section while converting the file into pdf.**