Text

Description automatically generated

Text

Description automatically generated

|  |  |
| --- | --- |
| our Name | ABHITEJ |
| Student ID |  |
|  |  |
| Which option did you use for this assignment? (Glue or Spark) | GLUE |
| Month and Year of the data files | FEB,AMRCH,APRIL |
| Number of lines in the combined CSV file | 1491673 |
| Size of the combined CSV file | 56.4MB |
| For Option #1 |  |
| Size of the **Parquet**file with **snappy**compression | 1.6 |
| Size of the ORC file with **snappy**compression | 1.3 |
| Size of the **Parquet**file with **gzip**compression | 1.1 |

Graphical user interface, application

Description automatically generated

Table

Description automatically generated

**Graphical user interface, text, application, email

Description automatically generated**

**Graphical user interface, text, application, email

Description automatically generated**

**Graphical user interface, application, Word

Description automatically generated**

**Graphical user interface, text, application, email

Description automatically generated**

**Graphical user interface, text, application, email

Description automatically generated**

**Text

Description automatically generated**

**Parqet zip has low size compare to other**