**Possible research questions and exploration by CSV file:**

Procedures.csv 🡪 Regression

-correlation of mortality rate with certain procedures

-is it statist. Significant?

-procedures: invasive line, intubation/extubation

A-line.csv 🡪 Classification

-which features upon admission (first day in ICU) are most closely associated with

Mortality?

-feature list:

-mortality day, SOFA, mean arterial pressure (MAP), heart rate (HR),

temperature, BUN, Creatinine, Chloride (Cl), potassium (K), hemoglobin (Hgb), Sodium (Na), white blood cell count (WBC)

Diagnosis\_ICD.csv 🡪 basic stats, visualizations, regression

-Top 10 ICU codes

-correlate ICD codes to mortality

-# of unique codes

DRGCODES.csv 🡪 (same as above)

ICUSTAYS.csv 🡪 basic stats (data.describe) filter on ‘LOS’ (length of stay)

Prescriptions.csv 🡪 basic stats (data.describe) filter on ‘Drug’

Patients.csv 🡪 basic stats filter on ‘Expired’

Need to mung data such that new CSV is created, OR we can run SQL queries instead. Ideally, we can look at following features:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Subject\_ID | A-Line features | LOS | Expired | Drug | ICD code | Procedures |
|  |  |  |  |  |  |  |