2) Perform the 5 data visualization tasks (given below) on the student performance dataset given in the link below. Each figure: 800×600 px, 300 DPI, title, labeled axes/units, legend if applicable, readable ticks. For each, add a 5–8 sentence interpretation to `reports. Before the visualizations (analysis step) you need to perform ingestion stage and preprocessing step (missing values etc.)` (10 points).

A. V1 — Gender boxplots (math vs reading) *(2 pts)*
   a. Question: Are there gender differences in math vs reading?
   b. Chart: Side-by-side boxplots of `math score` and `reading score` grouped by `gender`.

B. V2 — Test prep impact on math *(2 pts)*
   a. Question: Do students who completed test prep score higher in math?
   b. Chart: Any chart of your choice for `math score` by `test preparation course` (completed vs none).

C. V3 — Lunch type and average performance *(2 pts)*
   a. Question: Does lunch type (standard vs free/reduced) relate to outcomes?
   b. Chart: Grouped bar chart of mean `overall_avg` of all the scores (math, reading, writing) by `lunch`.

D. V4 — Subject correlations *(2 pts)*
   a. Question: How strongly do the three subjects move together?
   b. Chart: Correlation heatmap for `math`, `reading`, `writing` with annotated coefficients.

E. V5 — Math vs reading with trend lines by test prep *(2 pts)*
   a. Question: How strongly are math and reading scores associated, and do students who completed the test-preparation course have a *different slope* in the math–reading relationship than those who did not?
   b. Chart: Scatter plot with two straight best-fit lines (one for each group: `completed`, none).
      i. X-axis: `reading score`
      ii. Y-axis: `math score`
   c. Color: Points colored by `test preparation course` (legend must show the two groups and each group's n).

Data link: https://app.box.com/s/ji910ez3ycw137rw07xnhielxey7ww41

Submission:

Create a public GitHub repo and upload the folders for both the questions on the GitHub and submit the link to Canvas.