**GCD Assignments week 2**

*Activity 1*

Fill out the survey on: <https://docs.google.com/forms/d/1nyOZsOKbDIX559mmQgLw5i68rImEzZLxUe6tXgchh_0/viewform?usp=send_form>

After filling out the survey by all students, the results can be found on:

<https://docs.google.com/spreadsheets/d/1E6PlCvPHTN3Ngit0J46pcPAF1kKnVDB66F0xCThLb0w>

*Activity 2*

Inspect the data and find Quality measure violations

*Activity 3*

Give at least 5 scripts to solve violations, including a short statement why this is the best solution. It suffices to describe the scripts in words, like ‘delete all lines where attribute x contains non-text values’.

1) Column I: Remove non-numbers

2) Column J: Regex digit capture

3) Column C: Entries must contain a space and one token out of a list of known brands, replace unknown with other

4) Column E : Amount of entries must be bigger then a set number

5) Column B: If entry is outside STD discard

*Activity 4*

Which Missing Data Mechanisms (MCAR, MAR, NMAR) do you encounter? Explain.