

Project - 2:

Marked by: Aidan

Name: Ana Cardoso Cachopo

Points: 16/17

Acquire: Import Data Using Pandas

2/2 Good, and clear discussion

Parse: Perform EDA Using Pandas

2/2 Good, sensible exploration and cleaning

Mine: Plot Data & Generate Visuals

2/2 Very nice and good commentary. I think your weeks in the chart graph does not start its horizontal axis from zero, which is a bit misleading to the viewer. I liked the graph of ranking over time for one track - with something like tableau you could have a filter to show this graph for any arbitrary track quite easily.

Mine: Determine Correlations

0/2 Many people missed this so no worries, but it was part of the readme file and marking rubric so do check what is expected for each project rather than just the starter code which is incomplete. In any case you know better how to do this after week 3, but it could be done eg in seaborn also.

Identify: Problem Statement / Hypothesis

2/2 Ok, some good ideas here, but try to state it in a clear scientific way with a testing method, eg genre affects time in top 100 etc etc.

Identify: Risks & Assumptions

½ This is in the context of a clear hypothesis, what assumptions about the data could undermine it if false? You had good discussion about issues with the data during cleaning.

Refine: Evaluate Findings

½ You had some discussion in the visualisations but try to have a clear conclusion, and second guess it by discussing any issues. Also give next steps. Soon you will have better metrics also to justify your conclusions so you will be able to include that too.

Present: Describe Results in Blog Post

3/3 You didn't link this in the notebook but I remember we went through it so I give you the marks for that.

Bonus! Present: Write a short White Paper

3/3 Wow, you are the first who did this so far! Very nice!