Amazon Reviews

COM5508 Media Data Analytics

Lab 3

In this exercise, we will repeat last week's exercise of analysing Amazon reviews, now using pandas rather than plain python.

Analysing Amazon Reviews with Pandas

- Load the same data you used last week and read it into pandas data structures

- Use pandas functions to find/combine/calculate the same data you used last week to answer last week's research questions.

- Answer the research question by providing and interpreting the same key numbers or tables.

- If any of the numbers are different, explain why. Is there a mistake in last weeks code? Are you computing something different?

- Bonus: pick part of your analysis last week that you think is much easier (or much harder) in pandas than in plain python. Compare the two approaches qualitatively (comparing readability, compactness) and quantitatively (using %timeit). Which is better?

- Write the report in ipython-notebook, aim for about 2 pages, or 3 pages if you do the 'bonus' comparison.

- Submit in canvas as html or pdf