anacision Coding Challenge

Thank you for your application at anacision. To give you an insight into our work and to make sure we are a good match, we ask you to complete the following tasks and present them to our team next week. Plan to spend about 1-2 hours on each task. We prefer Python as the language of choice.

Task 1: CSV Viewer (Software Engineering)

Please write a robust, maintainable, and extensible command line application to display CSV files. Displaying the data in the terminal is sufficient, no GUI is necessary. Show us your implementation next week and walk us through your approach.

The logic is simple: it calls a CSV file, which should then be displayed in the terminal. When called, the file path should be specified.

The display should meet the following requirements:

- Each page is displayed with heading
- Each page consists of 10 data entries (rows)
- There is a cell marking (Exactly like in the example)
- The columns have a fixed width, which corresponds to the longest entry per column
- By entering the keys F/P/N/L/E, the respective commands can be executed (For this you could use the input() function)

The output should look like this

```
Name |Age|City    |
-----+---+--------+
Peter|42 |New York|
Paul |57 |London  |
Mary |35 |Munich  |
F)irst page, P)revious page, N)ext page, L)ast page, E)xit
```

The individual pages are represented as a table with headers and cell markers. It is assumed that the CSV files look like this:

- The first row are the headings
- The columns are separated by ";", the encoding is utf-8
- There are no line breaks

With the task you got an example of a file. Feel free to use it for implementation and presentation.

Task 2: Data Exploration

Please analyze the following dataset that contains information about flights and present us your approach next week: https://owncloud.exxeta.com/index.php/s/vBb4jCCMHopv5XR (Password: flights)

We also want to discuss the following questions: What could the information about drivers and the model be used for? What could a possible application look like and what would it take?

1. Perform an exploratory data analysis, especially regarding the question Which factors/drivers influence delays in arrivals (arr_delay) and departures (dep_delay)?
2. Build a predictive model for the binary variable of whether a **departure** is more than 30 minutes late (without using **arrival** delay). Use a suitable evaluation dataset for this purpose.