Exercise for Anne Lewis Strategies - Senior Data Scientist position

This repository has answers for each of the 4 sections:

- **ETL Exercise**: This section contains 4 deliverables:
  - *ETL_Code.ipynb*: A python code instance with documents. Running it requires installing python 3, pandas and jupyter notebook. *ipynb_checkpoints* folder should be disregarded. To run the file, please follow the subsequent instructions:
  - Open terminal
  - `pip install pandas`
  - `jupyter notebook`
  - Navigate to the *ETL_Code.ipynb* file and click on it
  - *people.csv*: A Comma Separated File that contains the people output dataframe
  - *acquisitions_facts.csv*: A Comma Separated File that contains the aquisition facts output dataframe
  - *ETL_email.docx*: A Word Document containing an email to the client as a follow-up from the ETL Exercise
- **Analytics Exercise**: This section contains 2 deliverables:
  - *Analytics_Data.xlsx*: A Excel File that contains 3 emails statistics
  - *Analytics_Answers.pdf*: A PDF file that contains answers to each of the 5 Analytical questions
- **Segmentation Exercise**: This section contains a unique deliverable:
  - *Segmentation_Answers.pdf*: A PDF file that contains an answer to each of the 3 Segmentation Quality Assurance questions
- **Proofing Exercise**: This section contains a unique deliverable:
  - *COMPLETED - Proofing Exercise.docx*: A Word Document that contains tracked changes of the proofing exercise