תרגיל 2 להגשה – תשפ"ב

ההגשה בהתאם לכל דרישות הקורס (סילבוס הקורס)

A.Create the following dataset (csv format) – dataset A -

1. Age (integer) - random number 16 - 70 (extra challenge: normal distribution)
2. Occupation (integer) – random number 1 - 20
3. purchase 1 (integer) - random number 0 - 5,000
4. purchase 2 (integer) - random number 0 - 10,000
5. purchase 3 (integer) - random number 0 - 2,500
6. purchase 4 (integer) - random number 0 - 25,000
7. purchase 5 (integer) - random number 0 - 15,000

Important notes:

1.Each record represent a buyer 2.The dataset must contain at least 20K records 3.The dataset/values must be personal per submission student

Submit via email:

1. **Full csv input file (based on the requirements @item A)**
2. **Code of all tasks + outcomes (per each task) –text/csv format**
3. **prepare to present /code and answers all tasks + the Open questions during the submission [see below 8.I & 8.III)**

B. Tasks – for each item (1-7: add the code and outcomes)

1. Find the average of the purchases (1 – till 5)
2. How many buyers between age 20-22 AND 30-34 AND 47-50
3. What AGE has the max average purchases (1 – till 5)
4. What is the purchase averages (1 till 5) per Occupation?
5. What are the min & max purchases, per Occupation per age?
6. how many buyers between ages 40-47 buy @purchase3 more than 2,200 and purchase4 less than 1,000 and purchase5 between 10,000-11,500?
7. What is the max purchase (1 till 5) per age?
8. Open question (**personal challenge**): based on **dataset A**, find a scenario with supported query/ies and AI model[s] that can contribute to a organization progress.

Note: free to add attributes, records, dataset[s] etc.

I. describe the full scenario (in the file – up to 5 lines)

II. run and find the full query/queries (with answers) that support the full scenario – (add the code and the outcomes)

III. describe the use of the outcomes [in 8.II] to the business progress based on 8.I scenario (in the file – up to 5 lines)