Quinn Nash

[Qnash@cnm.edu](mailto:Qnash@cnm.edu)

CIS 1275 C++ Programming Program 7 Part 1

**1. What is the website you used to obtain your data? List the agency/organization as well as the specific URL. Your instructor should be able to go and find your data from this information.**

Data found at: <https://www.kaggle.com/datasets>

URl: <https://www.kaggle.com/datasets/lamiatabassum/top-50-us-tech-companies-2022-2023-dataset>

**2. Often the data files contain more data than you’ll need for P7. What data from the file are you going to use/analyze in your P7?**

The data has one CSV file called Top 50 US Tech Companies 2022-2023. The data has 50 rows and 10 columns. The key columns I will use are:

* Employee Size (Column J/9) as integer
* Annual Income Tax (Column I/8) in billions as float
* Market Cap (Column G/6) in trillions as float
* Annual Revenue 2022-2023 (Column F/5) in billions as float
* Founding Year (Column E/4) as integer
* Sector (Column C/2) as string
* Company Name (Column A/0) as string

**3. You will need to perform analysis of the data and generate at least three data facts. Describe the analysis you plan to do.**

Find the company from each sector that has the largest employee size.

* Display company name, founding year, sector, and employee size.
* Also show which sector has the most employees and which sector has the most companies.
  + Are they the same sector?

Sum the income tax generated by each sector. Show the company from each sector that paid the highest ratio of income tax divided by revenue.

* Also show the company from each sector that paid the lowest ratio of income tax divided by revenue. (Don’t include companies who paid negative income tax)
* If a company paid a negative income tax, create a separate list that shows the company. Then find the average ratio of income tax divided by revenue for all companies and use it to suggest what the company should have paid in income tax.

Find the company from each sector with the highest ratio of Annual Revenue divided Market Cap

**4. What is your plan for reading the data? Include code to show your plan!**

Place csv file in same folder as P7, name it something simple like TechData.csv. I want a struct array to file information into for each company, 50 companies so size 50 array.

* **Struct TechCompanyInfo**
  + **String name “”**
  + **String sector “”**
  + **Int yearFounded 0**
  + **Float annualRevBillions 0.0**
  + **Float marketCapTrilllions 0.0**
  + **Float incomeTaxBillions 0.0**
  + **Int numEmployees 0**

Create ifstream object to read the file. Create a string for the line that is read in.

* **Ifstream fileinTechData;**
* **String lineInfo, String parseLineInfo**
* **Int row = 0, col =0**
* **TechCompanyInfo[50];**

Open and read file, check good open

* **FileTechData.open(“TechData.csv”);** //could make a filename constant
* **If (fileTechData.is\_open())**
  + **Cout << “Opened properly”**
  + **Getline(fileinTechData, lineInfo);** //clear header line first
  + **While(!fileinTechData.eof( ))**
    - * **Getline(fileinTechData, lineInfo);**
      * //create string stream to deliminate or parse through getline
      * **Stringstream ssLineInfor(lineInfo);**
        + **While (Getline(ssLineInfom, parseLineInfo, ‘,’));**
        + **If col == 0**

**TechCompanyInfo[row].name**

* + - * + **Else if col == 2**

**TechCompanyInfo[row].sector**

* + - * + **Else if col == 4**

**TechCompanyInfo[row].yearFounded**

* + - * + **Else if col == 5**

**TechCompanyInfo[row].annualRevBillions**

* + - * + **Else if col == 6**

**TechCompanyInfo[row].marketCapTrillions**

* + - * + **Else if col == 8**

**TechCompanyInfo[row].incomeTaxBillions**

* + - * + **Else if col == 9**

**TechCompanyInfo[row].numEmployees**

* + - * + **Col++** //increment columns
      * **Row++** //increment rows to change index of struct TechCompanyInfo
    - **FileinTechData.close();**
  + **Else**
    - **No open**