

Application of C in Data Analytics

Dr. Siddhartha Ghosh¹, Akshat Agrawal², Ramu B³, Kishore Kumar S⁴, Tharun Reddy N⁵

^{1,2,3,4,5}Dept of CSE, Vidya Jyothi Institute of Technology, Aziz Nagar Gate, Hyderabad

¹ siddhartha@vjit.ac.in

² agrawal.akshat2000@gmail.com

³ ramubabinolla@gmail.com

³ kishorekumar9889@gmail.com

⁴ tharunreddynalabolu99@gmail.com

Abstract: Data Analytics is a recent buzz word which is creating new sensation in the market. Traditionally it is the technique to analyze and represent data in a visual format which helps human to understand the story that collectively data has to say. Most of the cases the analysis of data is done using tools like Tableau and Qlik and with less coding. Whereas the coding based data analytics is done using Python and R. This paper is an initiative to improve research at UG level in engineering colleges. It is a outcome of a mini project where a professor and 4 of his B.Tech. students has tried to put some light on the concept of Application of C in Data Analytics. As the basic learner of coding and other computer techniques when this students are miles away from using high level programming languages like Python. As at second year level as they are comfortable with C programming languages, this paper made an attempt to find whether C can be used for Data Analytics or not. After a thorough study of the subject it is found that a combination of C, files and C graphics will help to take the work in next level. Also it uses infographics and comes out with a significant outcome which says that even C can be used for Data Analysis and representation. This paper is a initiative to inculcate the culture of research in UG level.

Keywords : Analytics , C-Graphics , Infographics, UG Research

1. Introduction

Data Analytics: The process of analyzing information to search out trends and answer queries, projecting information in the form of graph and charts like visual effects is popularly known as data Analytics. However, data analytics includes several techniques with many various goals. The data analytics method has some key parts like data gathering to filtering to analyzing to projecting. By combining these parts, a made information analytics initiative can offer an excellent service for the

business benefit or benefit to a govt or to a project which has many statistics of past and present.

Generally, this method begins with descriptive analytics which is like the method of describing historical trends in information and then showing it in graphical way. The main target of data analytics is to answer the question “what happened?” This typically involves mensuration ancient indicators like return income on investment (ROI). Also the business analytics doesn't create predictions or directly inform choices. It focuses on summarizing information in an exceedingly meaning and descriptive manner for the benefit of the business.

The next essential part of information analytics is advanced analytics. This part of information science takes advantage of advanced tools to extract information, create predictions and see trends. These tools embrace classical statistics further more as machine learning. Machine learning technologies like neural networks, linguistic communication process, sentiment analysis and a lot of change advanced analytics. This data provides new insight from information.

The use of machine learning techniques, big data, and low-cost computing power has enabled the utilization of those techniques in several industries. Data Analytics is also not different from it. It has the best algorithms, best programming languages like Python, R and then best tools like Tableau and Qlik in the market.

Though there are many tools and programming languages in the market the basic learners of programming which is basically C programming, feels that working in the area of Data Analytics or Data Science is something too posh and some time impossible. To give them a confidence and find a way to make the most popular language C also as an application language in Data Analytics this initiative is taken to do a small research work on application of C Language in Data Analytics and the final result is

satisfactory. Along with C we have considered C graphics to bring the visual effect on analyzed data.

C-Graphics: C-graphics is nothing but graphics in c, which is done with the help of pre-defined header file <graphics.h>. By default normal C development environment doesn't have graphics header file, so additionally we have to install and set up <graphics.h> .

Infographics : It is a processed part of the data analytics, and it is well known as the information graphics. An infographic is a representation of information in a graphics format to make the data easily understandable in brief. It includes line graph, pie chart, bar graph,...etc. It makes the quick analysis of the provided information, this helps the people to understand what's the information contained in the provided data.

2. Techniques of getting Visualization through Data Analytics

Data, Big Data, Analytics or Analyzing, visual representations are happening today and making the business profit maximum. Till few years back people used to totally dependent on tools like Excel and Access and regular SQL, PL/SQL queries. Today with few mouse clicks not only just data analysis, also happening great visualization. To name two popular tools here are Tableau and Qlik. In the programming domain it is Python and R which got maximum popularity for data analysis and visualization.

Tableau: Tableau is one in all the quickest evolving Business Intelligence (BI) and data analyzing tool. For a data analyst Tableau is fast, easy to learn and great result to come out at the end – kind of tool. Tableau has different choices to represent information in several views, applying filters, drill downs or format, making sets, groups, generating trend lines and showing prediction. The best feature Tableau are :

- Data mixing
- Real time analysis
- Collaboration of knowledge

Qlik : Qlik is often rated by analyst firm Gartner mutually of the highest information visualization and business intelligence (BI) vendors within the market, beside Tableau and Microsoft. Qlik's primary product, QlikView, supports visual information discovery, self-service atomic number coverage, and therefore the development and sharing of knowledge dashboards. Qlik Sense, supports a lot of free-form analytics and permits users to make information and internet applications through API connections.

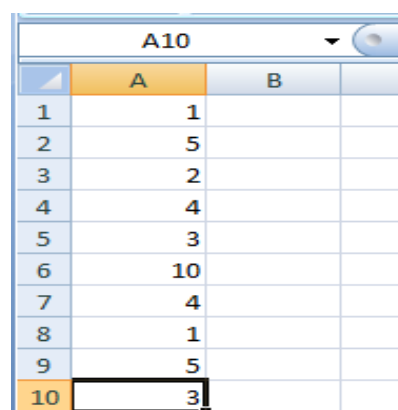
R : R is a language and atmosphere for applied mathematics computing and graphics. R provides a good sort of applied mathematics (linear and nonlinear modeling, classical applied mathematics tests, time-series analysis, classification, clump and graphical techniques, and is extremely protractible.

Python : Python is one of the latest languages which supports the depth and breadth of handling big data in the form of Data Analytics and Data Science. It does it through the different packages named as Numpy, SciPy, Matplotlib, NLTK, Theano etc. It has different packages to apply for different requirements in data handling and analyzing.

3. How to use C for Data Analysis in step by step ?

Gathering Data : The data which is to be analysed is either can be taken in *excel sheet(.xls)* or in *comma seperates values(.csv)*

Analyzing data : These excel sheets or csv files are called in the program using file operations in c (fopen() , fclose() etc..) .now this data is analyzed using simple mathematical formulae and arithmetic operations .



	A	B
1	1	
2		5
3		2
4		4
5		3
6		10
7		4
8		1
9		5
10		3

Fig1 : The Excel .CSV format and Data

Report : After analyzing of data , the analyzed data is represented in the form of line graph , pie chart , bar graph using c graphics .this form of representation of data (line graph , pie chart , bar graph) is termed as “info graphics” which is the part of data analytics .

Different functions of C graphics used for infographics are :

• **circle(int x, int y, int radius);**

Circle function is used to draw a circle with center (x,y) and third parameter specifies the radius of the circle.

• **line();**

line function is used to draw a line from a point(x1,y1) to point(x2,y2)

i.e. (x1,y1) and (x2,y2) are end points of the line.

• **initgraph();**

Which specify the exact position of representation.

• **initwindow();**

Which opens the new window .

• **SetColor();**

Which sets the colour to the given type ,it is based on enumerated data .x, y are coordinates of the point and third argument contains the address of string to be displayed.

4.Experimental Results :

```

-||-----<--!!-->>--- STATISTICAL DATA ---<<--!
1.LINE GRAPH
2.PIE CHART
3.BAR GRAPH

enter your option: 1

Read the file and store the lines into an array :
-----
Input the filename to be opened : linegraph.csv

The content of the file linegraph.csv are :
enter the marks in 4 subjects:
enter the marks in english:30
enter the marks in maths-2:45
enter the marks in pps:50
enter the marks in chemistry:20
    
```

Fig2 :



Fig 3:

Fig.2 & 3 shows the output screens for line graph. For the generation of line graph. First, the user enters the file name (with extension) of data whose line graph is required , then the data is processed and the resultant output is generated.

```

-||-----<--!!-->>--- STATISTICAL DATA ---<<--!
1.LINE GRAPH
2.PIE CHART
3.BAR GRAPH

enter your option: 3

Read the file and store the lines into an array :
-----
Input the filename to be opened : bargraph.csv

The content of the file 7 are :
ENTER THE NUMBER OF DATA ELEMENTS
    
```

Fig 4 :

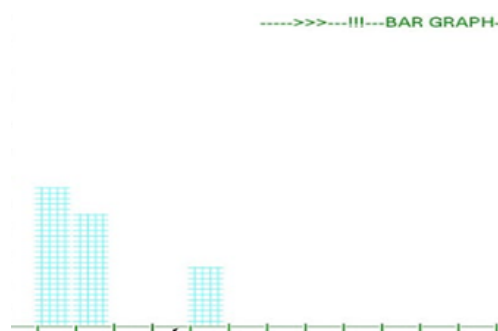


Fig 5 :

Fig.4 & 5 shows the output screens for bar graph for the generation of bar graph. first, the user enters the file name(with extension) of data whose bar graph is required, then the data is processed and the resultant output is generated .

```

-||-----<--!!-->>--- STATISTICAL DATA ---<<--!
1.LINE GRAPH
2.PIE CHART
3.BAR GRAPH

enter your option: 2

Read the file and store the lines into an array :
-----
Input the filename to be opened : pichart.csv

The content of the file pichart.csv are :

consider 60 student and enter
A. no.of students who got <=50%:5
B. no.of students who got between (50 to 60)% :20
C. no.of students who got between (60 to 75)%=:15
D. no.of students who got between (75 to 85)%=:10
E.no.of students who got between above 85%:10
    
```

Fig 6 :

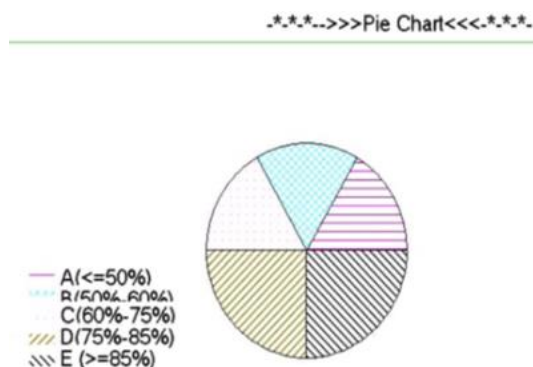


Fig 7 :

Fig.6 & 7 shows the output screens for pie chart .for the generation of pie chart. first, the user enters the file name(with extension) of data whose pie chart is required , then the data is processed and the resultant output is generated .

5. Conclusion

The working with C in the field of Data Analytics was a nice experience to understand the fundamental of data analytics. As a learner of C the team has always a wish to do something different in the application field of C. As to move into the modern areas like Data Science and Data Analytics with new tools and new languages will take time and proper planning, to gain more confidence this tiny research based work was taken in hand and the outcome is quiet interesting. This paper will surely help the c practitioners to fall more in love with C and its applications.

References :

- [1]ANDERSON F, PETERSC E, HALLORANJ, EVERYP,SHUTTLEWORTHJ,LIAROKAPISF, LANER, RICHARDSM.: In atthe deep end: An activity-led introduction to first year creative computing. Computer Graphics Forum 31, 6 (2012), 1852–1866
- [2] EYITAYOO T: Do students have the relevant ict skills they needto do their research projects. InProceedings of the 2011 Conference onInformation Technology Education(2011), SIGITE '11, pp. 287–292.
- [3] IHMSEN M, ORTHMANNJ SOLENTHALERB, KOLBA, TESCHNERM: SPH Fluids in Computer Graphics. InEurographics2014 - State of the Art Reports (STARs)(2014), pp. 21–42.

[4] JOELW. J.: Undergraduate research in computer science ed-ucation. InProceedings of the 18th ACM Conference on Innovationand Technology in Computer Science Education(2013), ITiCSE '13,pp. 361–361

[5] KOPPELMANH,VANDIJKB,VAN DERHOEVENG: Undergraduate research: A case study. In Proceedings of the 16th Annual Joint Conference on Innovation and Technology in Computer ScienceEducation(2011), ITiCSE '11, pp. 288–292