Sprint: Structured Approach to Data Analytics

Data Is Everywhere

With the upsurge of Big data and massive data growth, Data Analytics, as well as Data Science have become buzzwords. Since both Data Analytics and Data Science technologies are still emerging, the terms are sometimes used interchangeably, but the difference is stark. Let’s look at the definitions first:

Data Science

Is a combination of multiple disciplines – Mathematics, Statistics, Computer Science, Information Science, Machine Learning, and Artificial Intelligence. It includes concepts like data mining, data inference, predictive modeling, and ML algorithm development, to extract patterns from complex datasets and transform them into actionable business strategies.

Data Analytics

Is a combination of Statistics, Mathematics, and Statistical Analysis. It is the science of analyzing raw data to make conclusions about that information.

Data Analytics comes under the Data Science umbrella, with a focus on uncovering the specifics of extracted insights and fostering data-driven decision-making.

The first Sprint of the Data Analytics course is named A *Structured Approach to Data Analytics*. Another term, which is often mixed with Data Analytics, is Data Analysis. So what is the meaning of the other Data Analysis? Is it the same? The term definition is not set in stone and varies across the community. The Wikipedia provides a quite technical definition of data analysis:

Data Analysis

A process of inspecting, cleansing, transforming, and modeling data with the goal to discover useful information, inform conclusions, and support decision-making. Data analysis has multiple facets and approaches, encompassing diverse techniques under a variety of names, and is used in different business, science, and social science domains.

In the resources material of this course, Data Analysis is defined as:

The collection, transformation, and organization of data to conclude, make predictions, and drive informed decision-making.

All the definitions have in common that it is a process of examining/inspecting/transforming data to find something out in order to make informed/data driven decisions. According to the definition, data analysis is a part of data analytics and not a visa-versa.

The first Sprint will give you a general introduction to the data analyst profession. *The main data analyst responsibilities usually are:*

* To collect and interpret data.
* To identify relevant patterns in a dataset.
* To perform data querying using SQL.
* To experiment with different analytical tools like predictive analytics, prescriptive analytics, descriptive analytics, and diagnostic analytics.
* To use data visualization tools like Tableau, Looker, IBM Cognos Analytics, etc., for presenting the extracted information.

*The main data analyst skills usually are:*

* Well-versed in working with spreadsheets and SQL databases.
* Proficient in using tools like SAS, Tableau, Power BI, to name a few.
* Optionally proficient in R or Python programming.
* Adept in data visualization.
* Presentation skills to be able to share their finding convincingly.

The Data Analytics course material will help you to understand and obtain all of these necessary skills for a successful start of a data analyst career.

For this part, you should go through all the following materials. Get ready as if you're gearing up for an exciting Netflix binge, but this time, dive into the Introduction of Data Analytics. Be sure to take notes. You can write them on paper or use any other tool that you like. A good option is [Notion](https://www.notion.so/), which is simple to use and keeps all your notes in one place.

* [Data Analytics for Beginners](https://youtu.be/GbL-42kv5LI?si=6Po3_AKTEEgel5MT&t=336): (watch the entire video starting from 05:30 timestamp)
  + How Data Becomes Insights.
  + Working Within the Data Ecosystem.
  + How Data Informs Better Decisions.
* [Analytical Thinking](https://www.youtube.com/watch?v=CRVgnpda6Y8&list=PLTZYG7bZ1u6rcPlcX75pMkjsYY_HTL7aZ): (this video will cover)
  + Key Skills Data Analysts Possess.
  + Assessing Information Through an Analytical Lens.
  + Exploring Core Analytical Skills.
  + Using Data to Drive Successful Outcomes.
  + Real-world Data Magic.
* [Data analysis life cycle/phase](https://www.youtube.com/watch?v=dAyaThLvuug) (this video will cover)
  + The Data Life Cycle and Its Phases.
  + The Six Step Data Process.
  + Exploring Spreadsheets, Databases, and Visualization Tools.

[Six Steps of Data Analysis Process](https://www.geeksforgeeks.org/six-steps-of-data-analysis-process/) (another similar approach for the Analysis Process)

Take note that in the Turing College program, the main analysis process will be centered on the steps provided by Google:

1. Ask questions and define the problem.
2. Prepare data by collecting and storing the information.
3. Process data by cleaning and checking the information.
4. Analyze data to find patterns, relationships, and trends.
5. Share data with your audience.
6. Act on the data and use the analysis results.

* [Expanding on Google Six Steps](https://www.youtube.com/watch?v=Jye81eSts2c)

Analysis steps can vary across different organizations or structures. Below are some varied approaches presented:

* (Optional)[8 Steps in the Data LIfe Cycle - Harvard Business School](https://online.hbs.edu/blog/post/data-life-cycle)
* (Optional)[A Beginners Guide To The Data Analysis Process](https://www.youtube.com/watch?v=lgCNTuLBMK4)
* [Set up your toolbox](https://www.youtube.com/watch?v=N9Q3pOoVYIM&list=PLTZYG7bZ1u6rcPlcX75pMkjsYY_HTL7aZ) (You might skip this video, in case you already familiarized with the basic concepts of tools topics described below, but if you have time hop-in)
  + The Ins and Outs of Core Data Tools.
  + Spreadsheets demonstration (will be covered fully in Sprint 2).
  + SQL overview (will be covered fully in Sprint 3 and Sprint 4).
  + Data visualization overview.
* [The Power of Data in Business](https://www.youtube.com/watch?v=I189cilHQ9w&list=PLTZYG7bZ1u6rcPlcX75pMkjsYY_HTL7aZ) (this video will cover)
  + Data Analysts Across Industries.
  + The Power of Data in Business.
  + Understanding How Data Relates to Fairness.
  + Where do Data Analysts Work?

## **Additional Resource**

*Aditional resource part - This is for you if you feel the need to delve deeper into the details, but it's entirely optional and can be skipped. You can also return later to review the material at your convenience.*

* (Optional)[Daily life of a Data Analyst](https://www.youtube.com/watch?v=uSTtLpstV-o)YouTube is rich with diverse stories and perspectives in this area. feel free to watch one of the recommended videos, which provides a real-world glimpse into the routine and responsibilities of a Data Analyst.

Once you have gone through the given material, you are ready to complete this first Part! Click on the "I’m done" button above, which will lead to your first quiz. Once the quiz is complete and you’ve successfully finished this part, access the Part 2 — this is how you will progress through many of the Parts in Turing College.

Note that while some quiz questions might have a direct answer to them in the resources you’ve read. Use these questions as an indication of whether you are raising the right questions yourself and whether you are exploring them deeply enough.

If you don’t answer some of the questions correctly, do not worry – these quizzes are for self reflection only and will not impact your progress at Turing College in any way. Simply use them as a potential indicator for the topics that you can study further if you want to. If you disagree with an answer, first discuss it with your peers, JTLs or STLs and if you still think there’s an error in the answer, we would greatly appreciate it if you could message us via the support chat and let us know.

### **Good luck!**

# **QUIZ**

1.What is Data-driven decision-making?

A.The collection, transformation, and organization of data in order to draw conclusions, make predictions, and drive informed decision-making.

B.The way information is organized.

C.The various elements that interact with one another in order to produce, manage, store, organize, analyze, and share data.

D.Using facts to guide business strategy.

2.You may see the same quiz question repeated in further Parts in order to help you revise the concept and to encourage "spaced repetition"

A.True

B.False

3.Why is it not recommended to use a gut feeling in data analysis?

A.It is an intuitive understanding of something with little or no explanation.

B.It is not data-driven decision making.

C.Such decisions may be biased.

D.All of those are correct.

4.Which of the following is not a step in the data analysis process?

A.Prepare

B.Process

C.Share

D.Restore

5.Which is a step in the data lifecycle?

A.Analyze

B.Archive

C.Destroy

D.All of these are part of the data lifecycle

6.The main goal of a stand-up is to report on your progress, so that Turing College knows you are moving forward

A.True

B.False

7.Which of these best describes Analytical thinking?

A.Using facts to guide business strategy.

B.The management of the people, processes, and tools used in data analysis.

C.The various elements that interact with one another in order to produce, manage, store, organize, analyze, and share data.

D.The process of identifying and defining a problem, then solving it by using data in an organized, step-by-step manner.

8.A quiz makes up a part of the final score of a sprint, so a low score might prevent you from successfully passing the sprint

A.True

B.False

9.Which of the following is a common tool used to visualize data by a Data Analyst?

A.SQL

B.Tableau

C.VS code

D.Notepad

10.What is an important part of the Technical mindset?

A.The ability to break things down into smaller steps or pieces and work with them in an orderly and logical way.

B.The process of identifying and defining a problem, then solving it by using data in an organized, step-by-step manner.

C.The science of data

D.The various elements that interact with one another in order to produce, manage, store, organize, analyze, and share data.

11.What is Data Analysis?

A.The collection, transformation, and organization of data in order to draw conclusions, make predictions, and drive informed decision-making.

B.The way information is organised

C.The various elements that interact with one another in order to produce, manage, store, organize, analyze, and share data.

D.Using facts to guide business strategy.