Name of Skill

Data Analysis

Collecting and reviewing data, and making inferences from the data is called data analysis. Analyzing data is important in continuous improvement. Data allows you to make sound decisions about the process, product or service. So it is considered as a fundamental skills during software process improvement.

Classification of Skill

Data analysis can be classified as a Hard Technical Skill as requires training and experience to master. Hard skills are gained and improved by practise, training, and education.

Prerequisites for Skill

- Statistical Programming experience with R or Python
- Structured thinking approach and Critical Thinking
- Passion for solving problems
- Willingness to learn statistical concepts
- Love for numbers and quantitative stuff
- Grit to keep on learning
- Data Visualization
- Presentation Skills

Related Software Engineering Area(s)

- Software Engineering Management
- Software Engineering Models and Methods
- Software Quality

- Software Engineering Foundations

Rationale for Skill

Data analysis is an internal arrangement function done by data analysts through presenting numbers and figures to management. It involves a more detailed approach in recording, analyzing, disseminating and presenting data findings in a way that is easy to interpret and make decisions for the business. Data analysis help analysts to make decisions on process improvement and behavior prediction, increasing profitability and drive effective decision-making. Engineers can analyze the causes of particular events based on the data, understand the objectives and directives for your business, and you will have technical insights using an easy to understand language.

Roles for Skill

- Process Improvement Analyst
- Research Assistant
- Business Analyst
- Data Analyst
- Software Engineer
- Data Scientist

Work Related to Skill

- Data analysis, auditing, integration testing and system implementation reporting assistance received.
- Quality and service problems examined by root cause data review, analyses of waste materials and implementation of programs.
- Presented support for data processing using pivot tables, formulas and MS Excel graph / chart manipulation.

- Conduct data analysis, monitoring, study of implementation and troubleshooting.
- Conduct data analysis and reviews of process changes for practice sites to enhance performance, patient safety, patient satisfaction and staff satisfaction.

Real-World Example

As a real world scenario, to analyse issues that businesses have faced with their software process improvement (SPI) projects, we need to ask what are the company's software development challenges and concerns are, which are the obstacles inside the business for SPI? The huge amount of information produced from the practises of Software Company has complicated the extraction of knowledge and the SPI initiatives. Data analysis is a potential strategy for effective information management. However, using the best technique in data analysis can lead to impactful results, e.g. using process mining is a unique approach that uses data mining and business process modelling elements to improve processes.

Role of Academia or Industry in Cultivating the Skill

As demand for data analysis and data science continues to rise, academia are trying to find out how best to contribute to the training of a workforce. More businesses of all sizes and in all industries see data processing as a vital part of their success. Strong demand for data analytics graduates makes data science courses a prerequisite for all students in this area. On the other hand, the growth of data analysis in academia is primarily a function of its vast private sector expansion, and many of the major developments in data science over the past decade have been made by large private companies such as Google, Facebook and LinkedIn.

Tools Supporting the Skill

The list of top 5 of data analytics tools, both open source and paid version, based on their popularity, learning and performance.

- **R Programming:** the leading analytics tool used for statistics and data modeling. It can easily manipulate your data and present in different ways.
- **Tableau Public:** a free software that connects any data source and creates data visualizations, maps, dashboards.
- **Python:** an object-oriented scripting language which is easy to read, write, maintain and is a free open source tool.
- **SAS:** a programming environment and language for data manipulation and a leader in analytics
- **Apache Spark:** a fast large-scale data processing engine and executes applications in Hadoop clusters 100 times faster in memory and 10 times faster on disk.

Skill Self-Assessment

Although I have a good knowledge and experience about the pre-requisites and tools of this skill, being expertise in this area require minimum 1-2 years of quantitative analysis work experience and more than 2 years experience handling, manipulating and analyzing data and creating analytical reports. I try to combine this hard skill with soft skills to create a good balance and be a good team player.