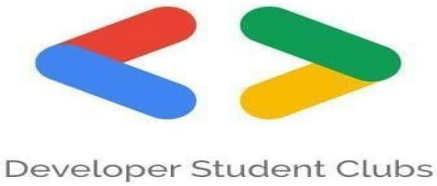

Google Developers

Developer Student Clubs

DSC – HTI

Data Science Track – Level 1

2020

DSC

Helping students bridge the gap between theory and practice, Developer Student Clubs (DSC) are university-based community groups for students interested in Google developer technologies. Students from all undergraduate or graduate programs with an interest in.

Data Science Track

The Data Science track is long-term training track for learning students the fundamentals and practices of how to process, manipulate and deal with data, in standardized and scientific manner.

In this level, we provide clear and very good training of how to deal with structured data, in methods of data understanding, preparation, preprocessing, visualization and basics of modeling.

Team

Hend Emad (Leader)

Omar Mohamed Mounir (Data Science Instructor)

Mohammed A. El-Agha (Data Science Instructor and Mentor)

Content

Stage 1

1. Data
2. Data Science
3. Python
4. Data Collecting and Understanding
5. Data Visualization
6. Data Preprocessing
    a. Data Cleaning

i. Missing Data

Stage 2

ii. Noisy Data
    1. Binning
    2. Regression
    3. Clustering
iii. Inconsistency

b. Data Transformation
    i. Type Conversion

Stage 3

ii. Normalization and Scaling
iii. Attribute Selection
iv. Discretization
v. Concept Hierarchy Generation

c. Data Reduction
    i. Attribute Subset Selection
    ii. Dimensionality Reduction
    iii. Sampling

Stage 4

7. Data Modeling
8. Projects and Exercises

Time Table of Content

Session #	Topics
1,2	Data, Data Science, Python
3	Data Collecting and Understanding
4	Data Visualization
5	Intro to Data Preprocessing, Missing Data
6	Noisy Data
7	Noisy Data, Inconsistency
8	Type Conversion
9	Normalization and Scaling
10	Attribute Selection
11	Discretization, Concept Hierarchy Generation
12	Attribute Subset Selection, Dimensionality Reduction
13	Sampling
14,15	Data Modeling (overview)
16,17, …	Projects and Exercises