

**Faculty of Computers and Artificial Intelligence
(Helwan University)**

**SW: Selected Topics in Software Engineering (Data
Science and Big Data)**

Task 1

1. Table 1 represents the collected data for a set of car models:

Table 1: Data that describe a set of car models

Car Model	Manufacturing Year	Cylinders	CC	Price (L.E.)	Maximum Speed (Km/hour)	Horse Power	Used/New
Toyota	2009		1600	150000	195	100	New
Jeep	2009	6	3700	320000	200	210	New
Mercury	2008	6	4000	2500000		210	
Opel	2008	4	1600	180000	192	105	New
Mitsubishi	2006	4	1600		170	106	Used
Honda	2009	4	1500	120000	180	92	New
Mazda	2003	4	1600	85000	180		Used

Apply the following points (In code):

- a. Fill-in the incomplete using mean or median or mode methods.
- b. Remove outlier in “Price” field.
- c. Convert Categorical columns to Numerical columns.