CIND 119
Introduction to Big Data Analytics
Lab 1

1. Download **intro_auto.csv** from your D2L course shell, under Course Materials, in the "Lab CSV Files" folder. Complete the following tasks.

   a. Read the file in SAS using the **import** procedure. Check the log window to confirm the file is loaded correctly.

   b. **Print** the contents of the file on the Results Viewer Window.

   c. Analyze the metadata of the file in SAS using the **contents** procedure.

2. Apply **means**, **freq** and **univariate** procedures to inspect statistical properties of the data.

3. Apply the **univariate** procedure to visualize the distribution of **mpg** (miles per gallon) and **price** columns of the data.

4. Download **intro_auto_noisy.csv** from your D2L course shell, under Course Materials, in the **"Lab CSV Files"** folder. Complete the following tasks:

   a. Write the SAS code to **import** and **print** contents of the file.

   b. Notice the missing values and duplicate rows (20, 21 and 26, 27). Remove the duplicate rows.

End of CIND119 Lab 1

Ceni Babaoglu, PhD