**Car Evaluation Data Set**

Baza de date auto de evaluare a fost derivată dintr-un simplu model de decizie ierarhic dezvoltat inițial pentru demonstrarea DEX.

1. ***A. Formularea problemei***
2. **B. *Metode și tehnici de soluționare a problemei* – noțiuni teoretice necesare abordării problemei enunțate (scurtă descriere).**
3. **C. *Experimente efectuate* – prezentarea etapelor pentru atingerea obiectivelor (rezolvarea cerințelor)**
4. **1.Încărcarea fișierului sursă – prezentare structura fișier (x variabile numerice, y variabile categoriale etc.).**

Primul pas efectuat pentru analiza Car Evaluation Data Set a fost download-area fisierelor sursa Data Folder/Data SetDescription din baza de date publică UCI Machine Learning Repository.

Derivată dintr-un model de decizie ierarhic simplu, această bază de date poate fi utilă pentru testarea inducției constructive și a metodelor de descoperire a structurii.

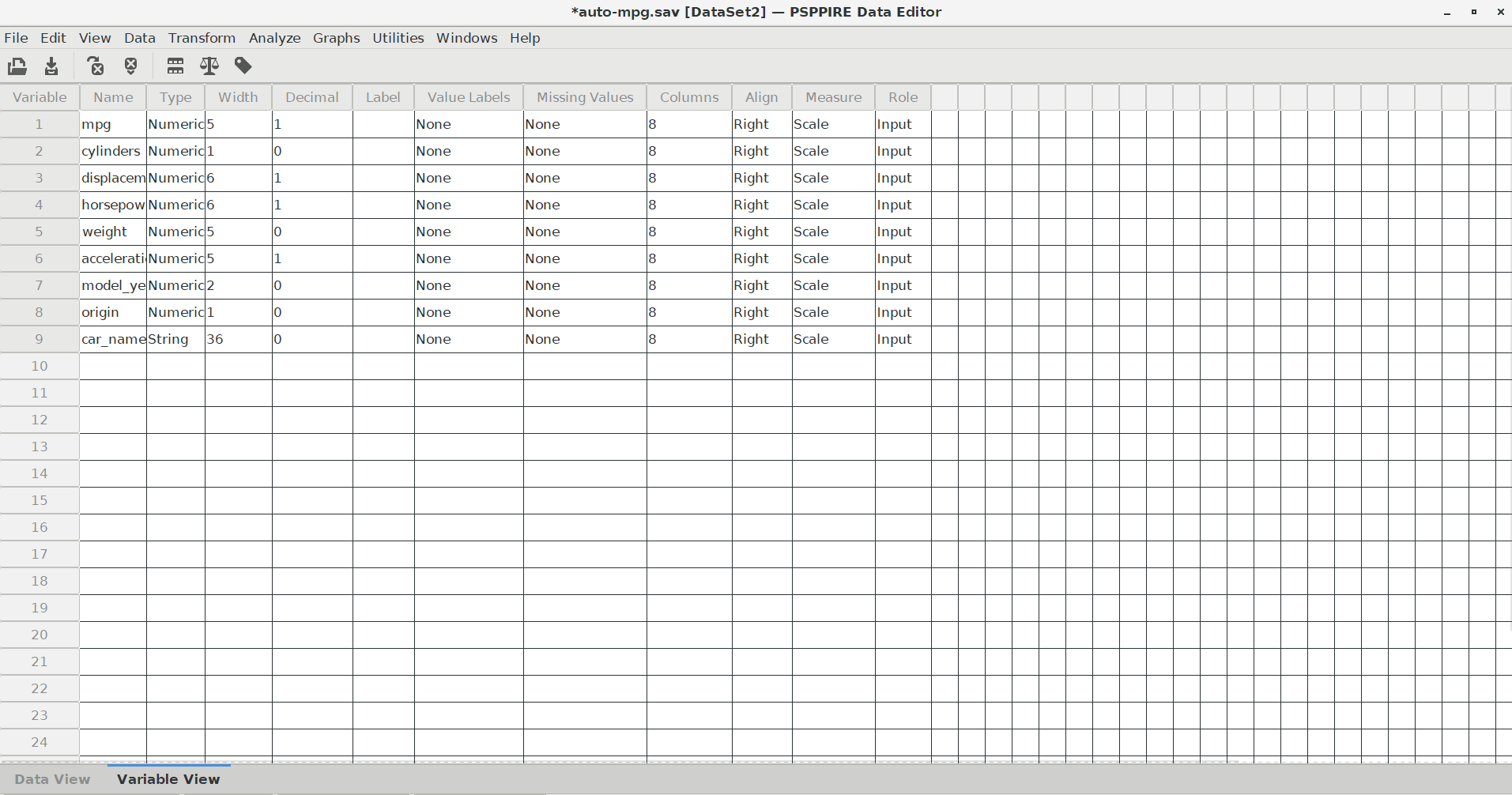
Fisierele sursa download-ate (car.c45-names, car.dat, car.names) inainte de a face orice fel de analiză statistică cu PSPP, au fost redactate cu ajutorul aplicatiei Notepad++

Din motiv ca a fost cu valori lipsă, a fost redactata si modalitatea de aliniere , doar apoi salvate ca fisiere cu extensia .sav si apoi importante in aplicatia grafica PSPP.

DATA VIEW:

A screenshot of a computer

Description automatically generated

1. VARIABLE VIEW:
2. 
3. **2.Analiza exploratorie a datelor - statistici descriptive ale variabilelor considerate.Discuție.**

Pentru a calcula valoarea minimă, valoarea maximă, media,

abaterea standard, asimetria pentru variabila (cylinders) a fost folosit Descriptive Statistics Frequencies.

A screenshot of a cell phone

Description automatically generated

În ambele tabele se regăsesc valorile variabilei (cylinders) , frecvențele absolute, procentul, procentul

cumulat, în timp ce al doilea tabel conține valorile indicatorilor statistici medie, abaterea

standard, asimetria, valoarea maximă, respectiv valoarea minimă.

Calcularea statisticilor descriptive: valoarea minimă, valoarea maximă, media, abaterea

standard, dispersia și asimetria pentru atributul (WEIGHT).

A screenshot of a cell phone

Description automatically generated

Reprezentarea grafic pentru variabila (WEIGHT), utilizând histograma:

A screenshot of a cell phone

Description automatically generated

Reprezentarea grafică sub fomă de bar chart pentru variabila (ACCELERATION):

A screenshot of a cell phone

Description automatically generated

Am realizat un grafic de tip scatter plot și am interpretaț rezultatele. Am droit sa verific daca există o legătură între anul de fabricatie a autoturismului (variabila: model\_year) si aceleratie (variabila: acceleration)

A screenshot of a cell phone

Description automatically generated

Dupa care a fost analizata corelația dintre variabila(displacement) ce reprezinta distanta parcursa si (mpg) care reprezinta consumul de combustibil . Astfel, între distanta de deplasare și consumul de carburant 1L/100km există o corelație de -0.80.

A screenshot of a cell phone

Description automatically generated

1. **3.Realizarea de grafice corespunzătoare variabilelor (minim 3 grafice) și interpretarea acestora.**
2. Pentru a furniza reprezentări grafice de date, cum ar fi histograme sau boxplot-uri se consideră fișierul sursă auto-mpg.sav.
3. **4.Verificarea corelației între variabilele considerate. Exemplu, interpretare rezultate.**
4. **5.Teste statistice pentru compararea mediilor – exemplu, interpretare rezultate.**
5. Se dorește a se verifica ipoteza următoare: Greutatea media a vehiculelor din lista este de 2600 kg (H0=2600). Intervalul de încredere considerat este 95%(default). Astfel, pentru ipoteza enunțată, se alege variabila (weight) și se precizează valoarea medie cu care se va compara media variabilei WEIGHT (H0=50).
6. A screenshot of a cell phone

   Description automatically generated
7. Rezultatele obținute indică o valoare p=0.000 și o valoare t=8.73 si df=397.
8. **6.Analiza variației (ANOVA) – exemplu, interpretare rezultate.**
9. **7.Analiza în componente principale – exemplu, interpretare rezultate.**
10. **Doar o singura componenta se vac rea din cele 8 variabile**
11. ***D. Concluzii generale***
12. ***E. Referințe bibliografice/webografie.***