



Adatelemzés **SAS Enterprise Guide használatával**

Soltész Gábor
[solteszgabeee\[at\]gmail.com](mailto:solteszgabeee[at]gmail.com)

›Tartalom

- SAS Enterprise Guide bemutatása
 - Kezelőfelület
- Adatbeolvasás
 - Szűrés, rendezés
 - Új változó létrehozása
- Elemzések készítése
 - Leíró statisztika készítése
 - Korreláció, regresszió számítás

> SAS Enterprise Guide

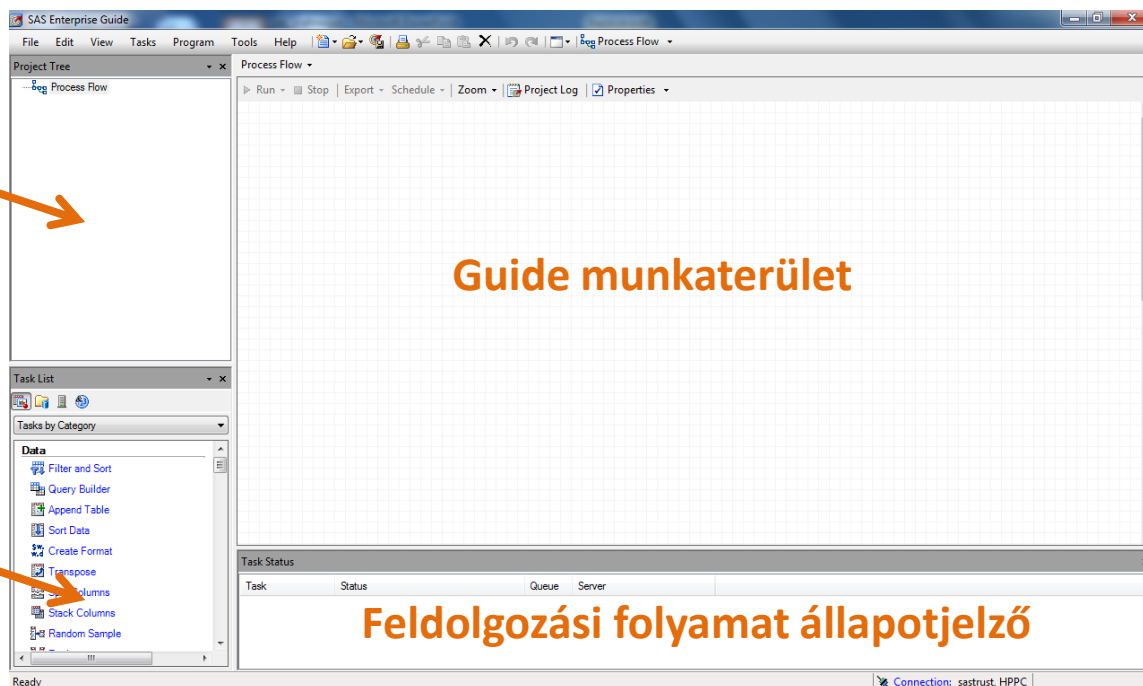
- Mire használható
 - Adatokból információ kinyerése
 - Listák, táblák készítése
 - Adatok vizuális megjelenítése (grafikonok)
 - Statisztikai elemzések készítése
 - OLAP kockák megjelenítése és kezelése

➤ Kezelői felület bemutatása

- SAS Enterprise Guide elindítása
 - Start menü / Programok / SAS / Enterprise Guide

- Felület

Project Explorer ablak



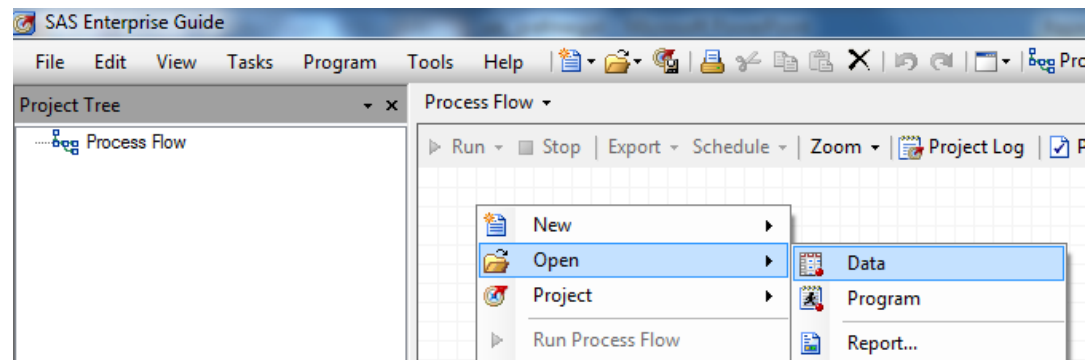
Folyamat lista elemek

Guide munkaterület

Feldolgozási folyamat állapotjelző

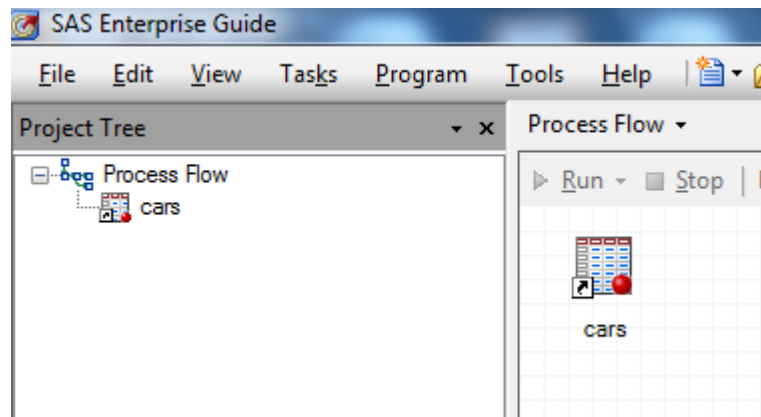
> Első lépések

- Új projekt létrehozása
 - File / New / Project
- Adattábla beolvasása
 - Jobb klikk a munkaterületen / Open / Data



>Autók tábla (cars) beolvasása

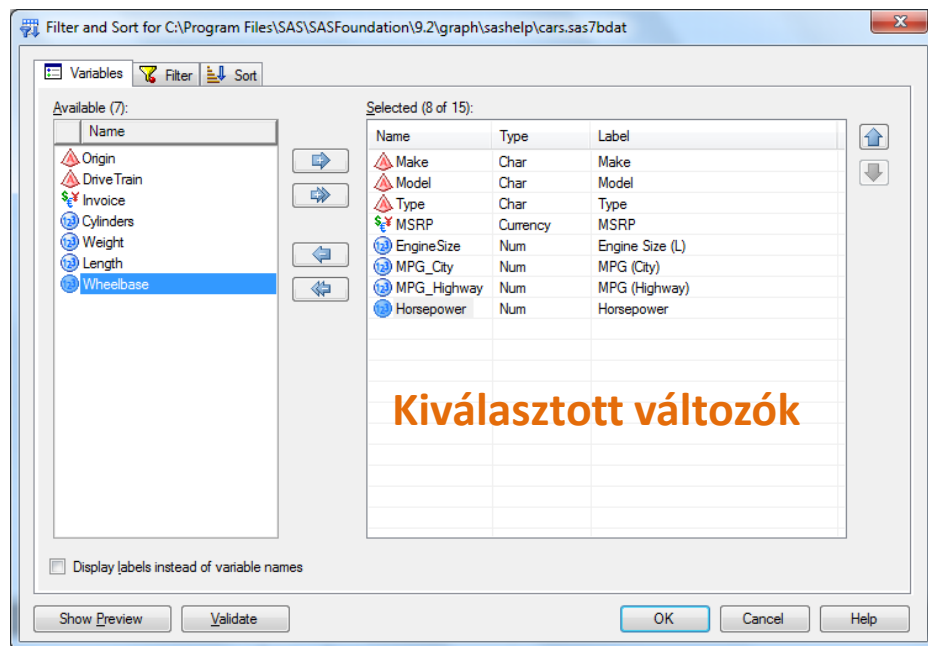
- Cars adattábla kiválasztása
 - *C:\Program Files\SAS\SASFoundation\9.2\graph\sashelp*
- Munkaterületen megjelenik a tábla



- Egyszerre több táblát is betölthetünk

> Változók kiválasztása

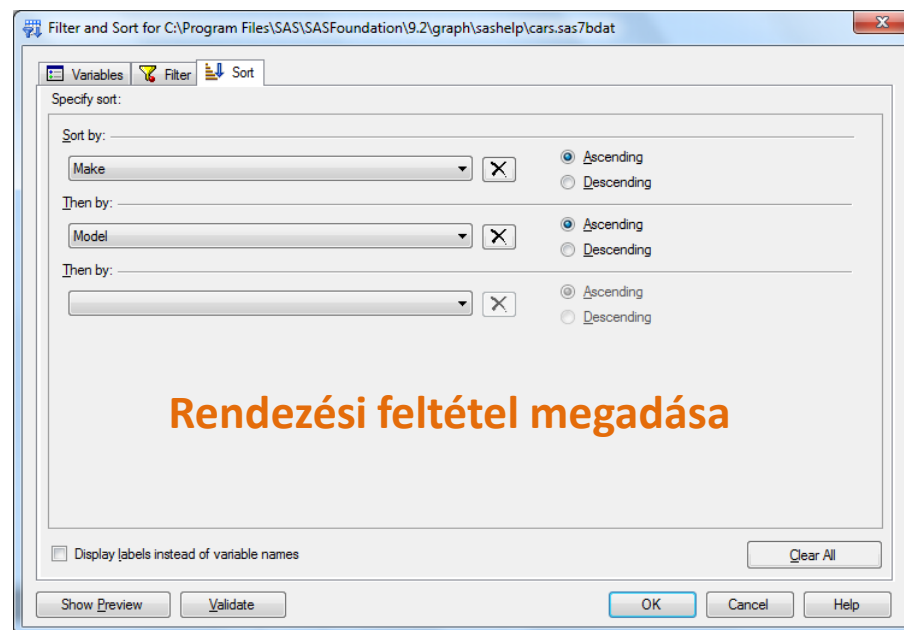
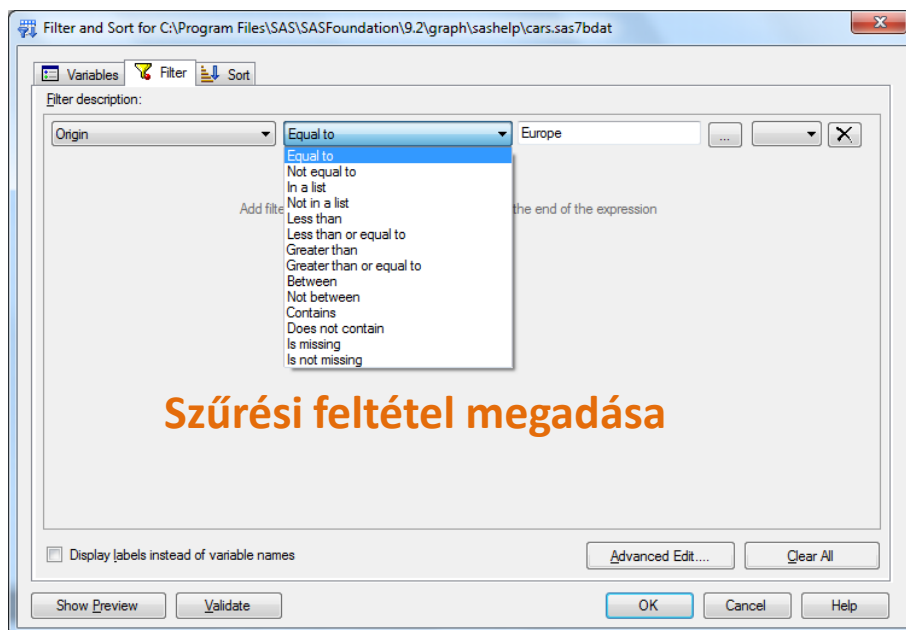
- Jobb klikk az adatállományon / Filter and Sort
- Változók listája típus szerint
 - Mennyiségi ismérv
(numerikus változó)
 - Minőségi ismérv
(osztályozó változó)



> Szűrések rendezés

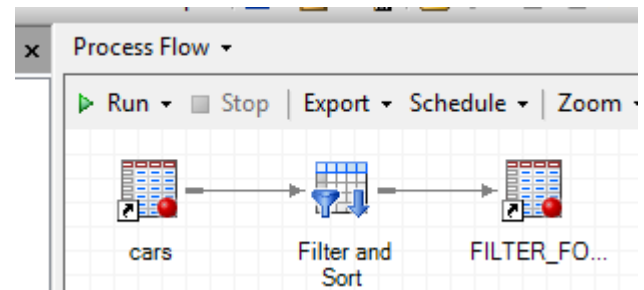
- Filter fül

- Sort fül



> Munkaterület

- Minden művelet végrehajtása után a folyamatábra változik



- Folyamat végrehajtási iránya balról jobbra
 - Esetünkben a második adattábla a mérvadó csak az abban definiált változók használhatók

› Új változó készítése

- A Guide alapértelmezés szerint csak olvasási jogokkal nyitja meg a beolvasott táblákat (read-only mode)
- Update mode: Edit / Protect Data
- Így már módosíthatjuk a tábla struktúráját
- Nyissuk meg az általunk leszűrt táblát
 - Jelöljük ki az utolsó oszlopot / Jobb klikk / Insert Column

> Új változó készítése/2

- Határozzuk meg az árat forintban

Változó neve

Kifejezés szerkesztő

Változó típusa

Insert the new column
 To the left To the right

Name: Price

Label: Ár forintban

Type: Numeric Length (in bytes): 8

Group: Numeric

Expression: Note: The expression will be applied only once and will not be stored.

Advanced Expression Builder

Enter an expression:
MSRP*200

Home Next Back End Undo Redo Edit Favorites

+ - * / ** || (x) 'x' "x" , 'abc'n

FILTER_FOR_CARS

- Make
- Model
- Type
- MSRP
- EngineSize
- MPG_City
- MPG_Highway

The maximum number of discrete values is limited to the first 100,000 rows of data.

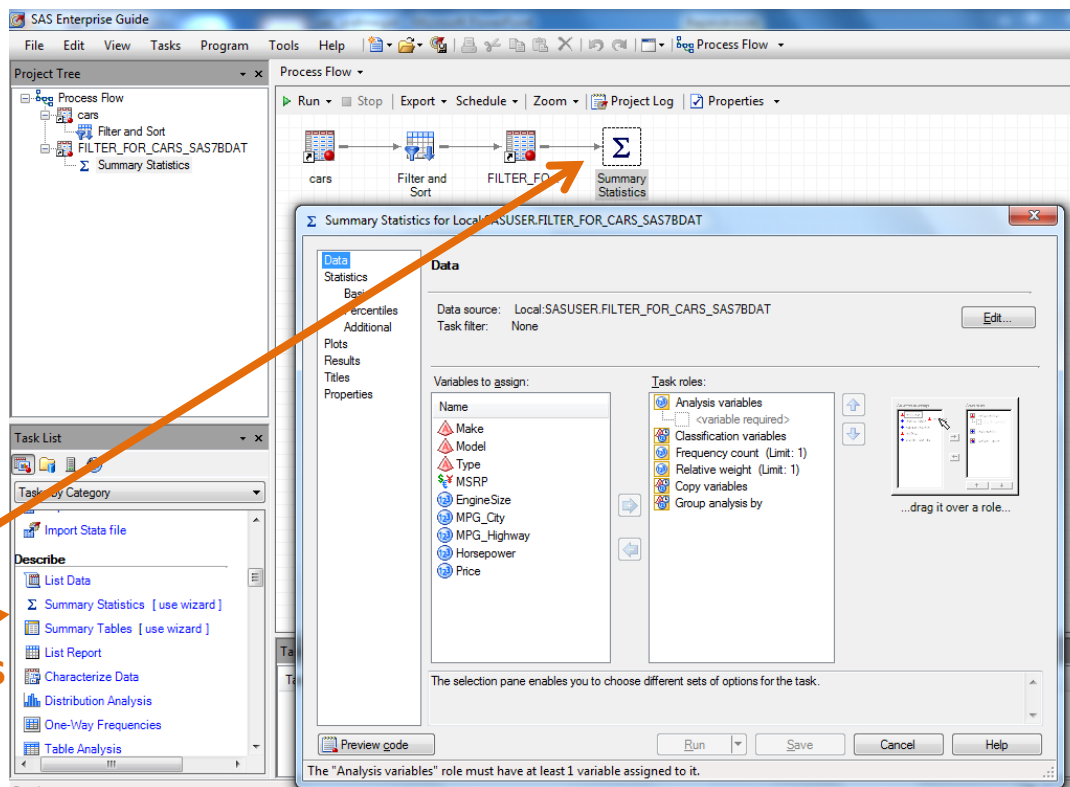
Get Values Select Values

OK Cancel Help

› Leíró statisztika készítése

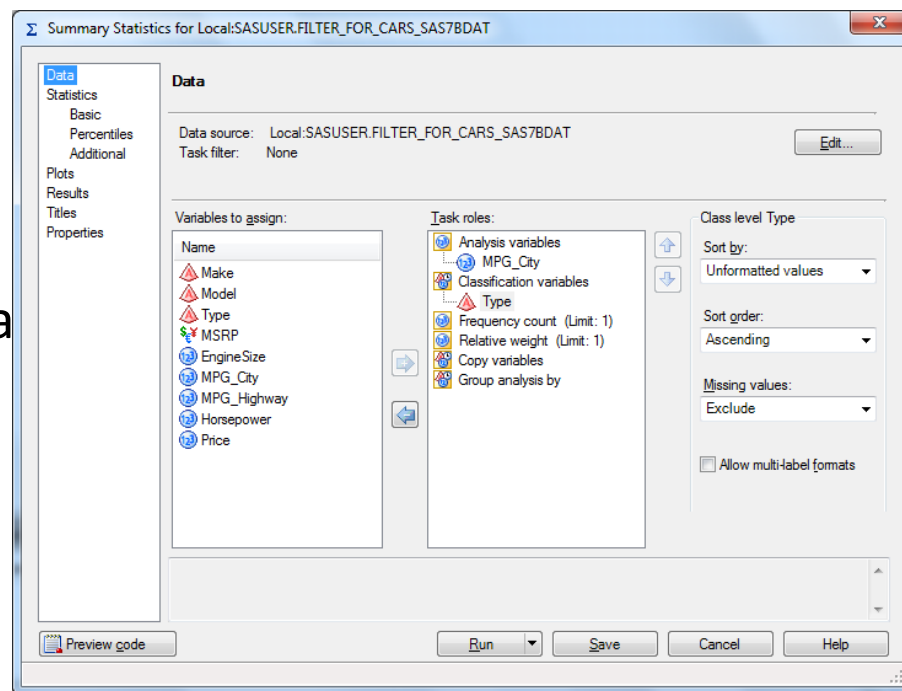
- Válasszuk a Summary Statistics listaelemet

Summary Statistics



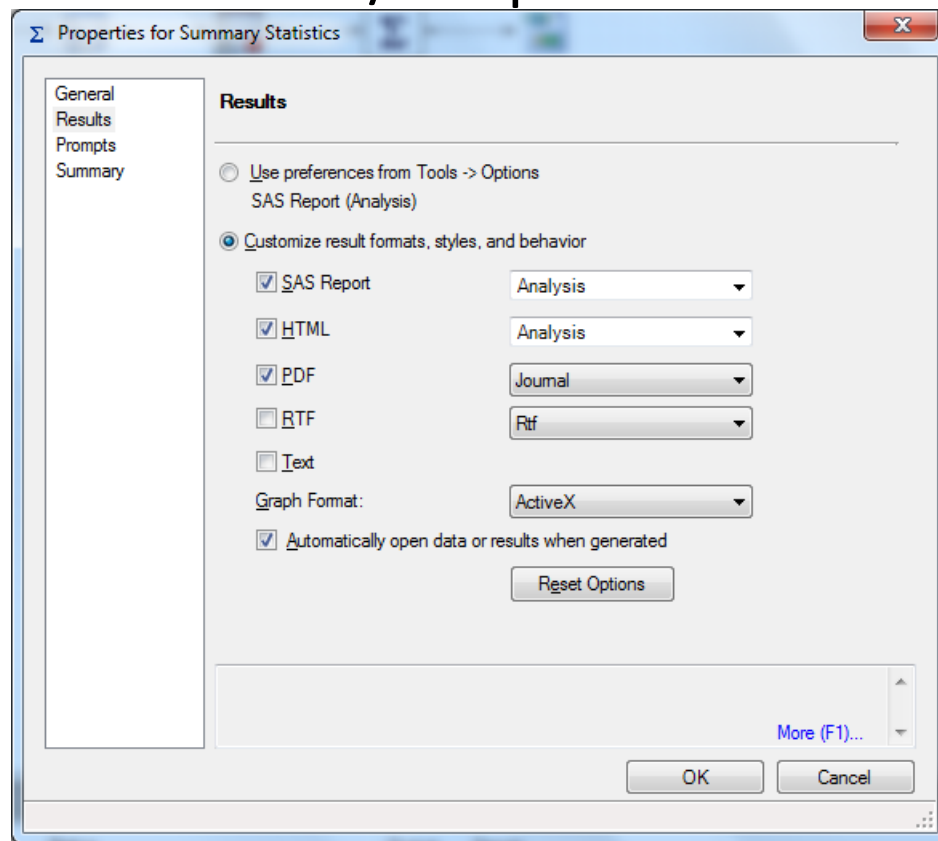
> Leíró statisztika beállítása

- Data menüpont
 - Elemzendő változó: MPG_City
 - Osztályozó változó: Type
- Statistics
 - Statisztikai változók kiválasztása
- Plot
 - Grafikonok kiválasztása
- Titles
 - Címkék megadása

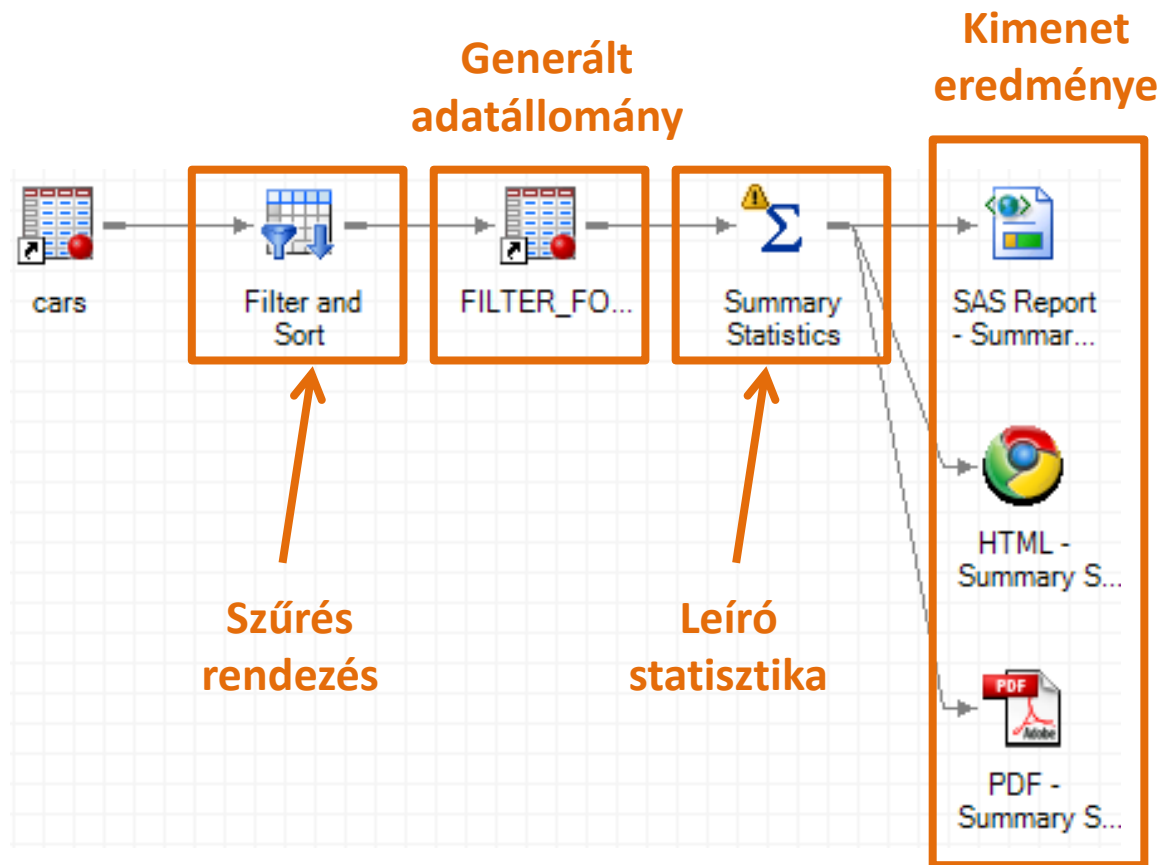


>Eredmény formátum módosítása

- Jobb klikk a Summary Statistics elemen / Properties
- Result menüpont
 - Válasszuk ki a formátumokat
 - HTML
 - SAS output
 - PDF
- Klikk az OK gombra
- Beállítás Véglegesítés
 - Jobb klikk az elemen / Run



>Folyamatábra

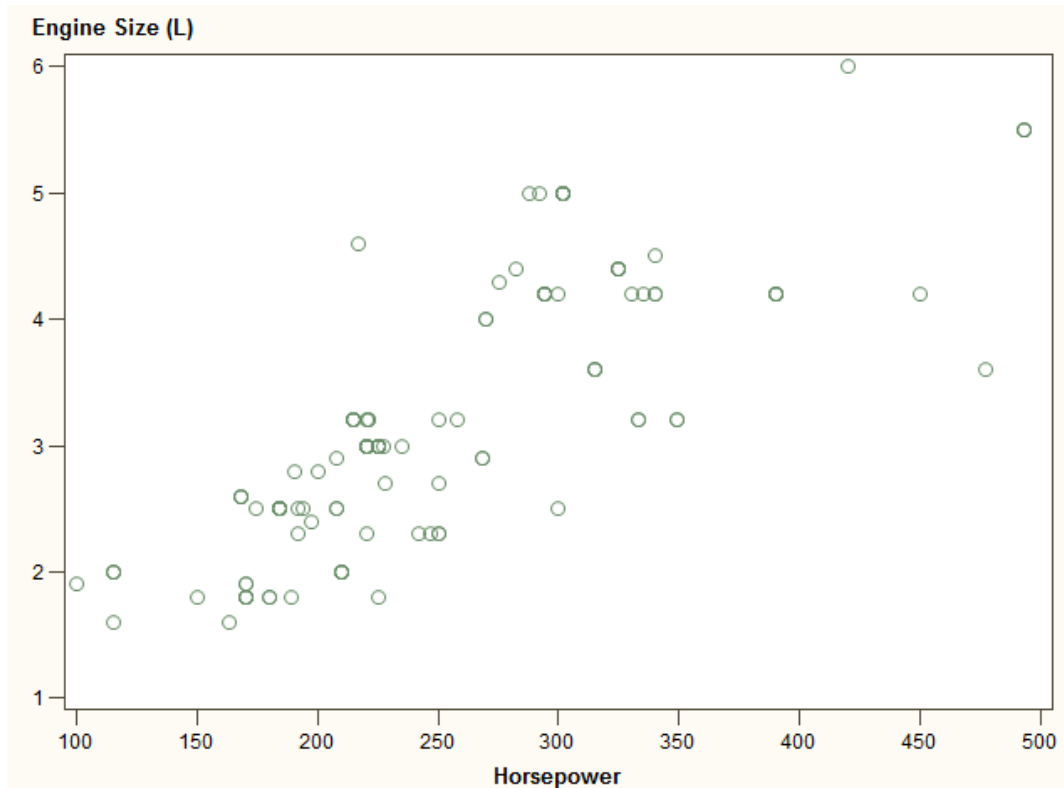


> Korreláció és regresszió számítás

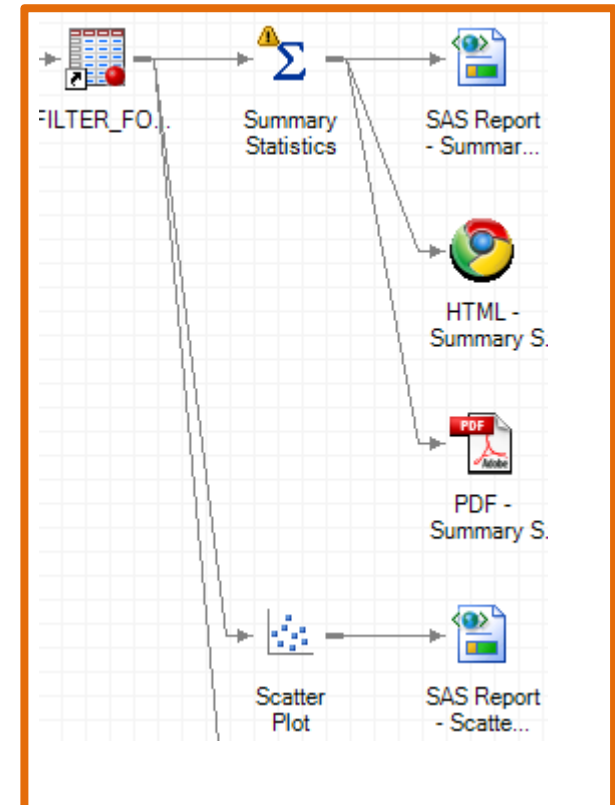
- Korreláció számítás
 - Mennyiségi ismérvek közötti sztochasztikus kapcsolat
 - Vizsgálja, hogy egy jelenség alakulását egy másik jelenség hatása befolyásolja e
 - Elemzés előtt célszerű pontdiagramot (Scatter plot) készíteni az elemzendő változókról
- Készítsünk pontdiagramot a motorméret (Engine size) és a teljesítmény (Horsepower) kapcsolatáról

> Pontdiagram

- Task menü / Graph / Scatter plot

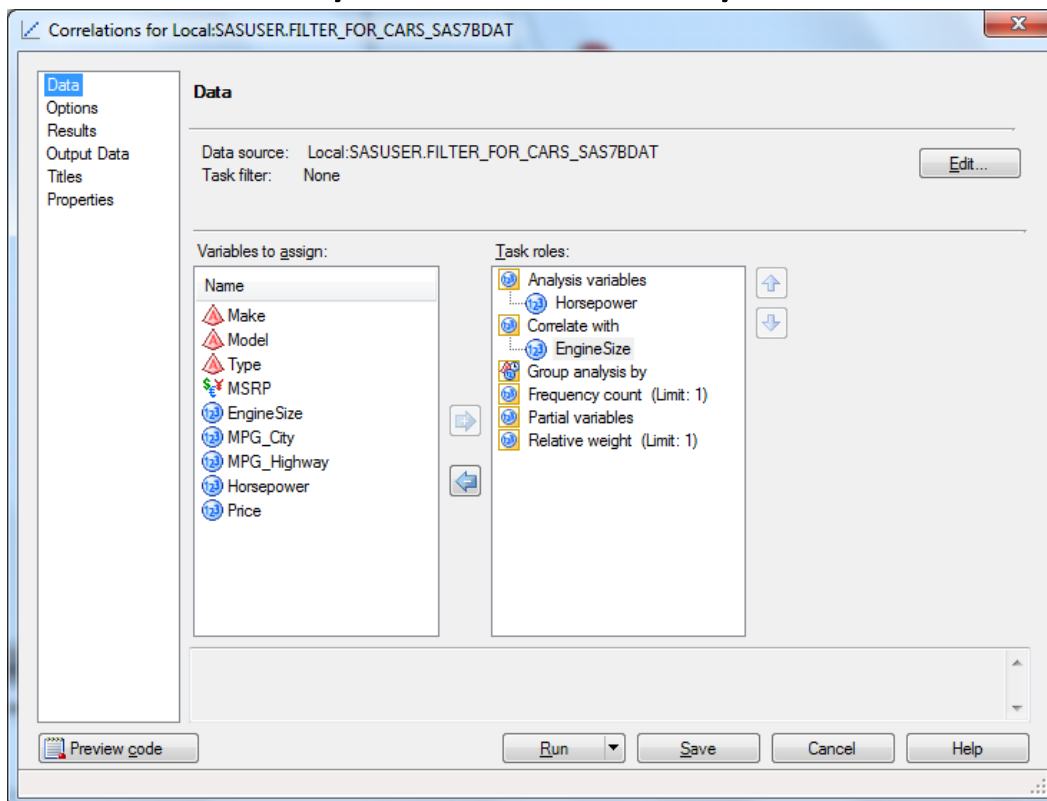


Folyamatábra

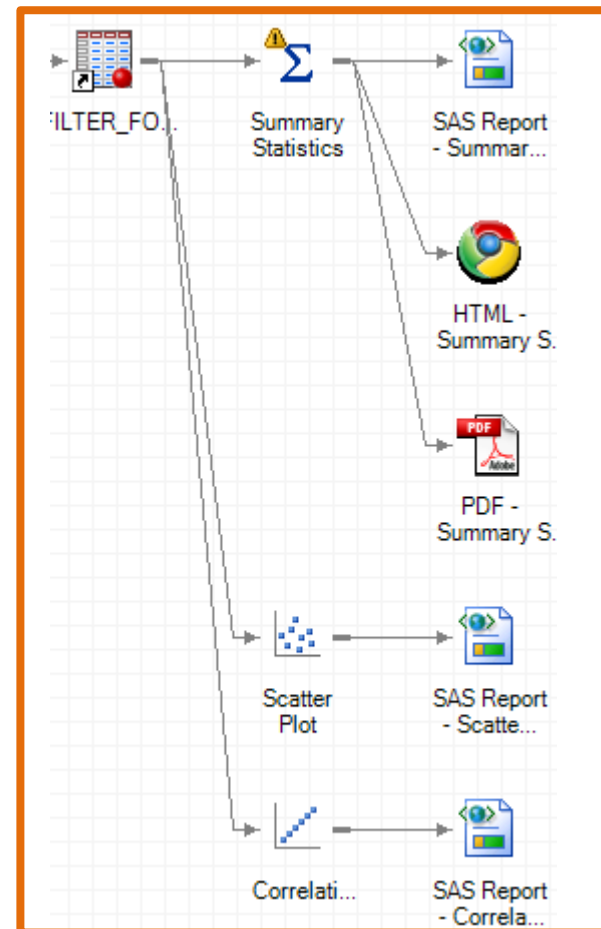


> Korreláció vizsgálat

- Task menü / Multivariate / Correlations



Folyamatábra



> Korreláció számítás eredménye

Correlation Analysis
The CORR Procedure

1 With Variables:	EngineSize
1 Variables:	Horsepower

Simple Statistics							
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum	Label
EngineSize	123	3.20650	1.03716	394.40000	1.60000	6.00000	Engine Size (L)
Horsepower	123	251.89431	80.73888	30983	100.00000	493.00000	

Pearson Correlation Coefficients, N = 123 Prob > r under H0: Rho=0	
	Horsepower
EngineSize	0.80036
Engine Size (L)	<.0001

R=0 nincs kapcsolat
R<0.5 gyenge kapcsolat
R<0.8 közepes kapcsolat
R >0.8 erős kapcsolat

- Eredmény értékelése
 - R=0.8 Erős a kapcsolat a motorméret és a teljesítmény között

> Regresszió számítás

- Regresszió számítás

- A korrelációval meghatározott kapcsolatnak a természetét írja le
- Feltételezzük, hogy lineáris a kapcsolat
- Egyenlet $\bar{Y} = B_0 + B_1 * X$
 - B_0 tengelymetszet
 - B_1 meredekség

- X: eredmény (függő változó)

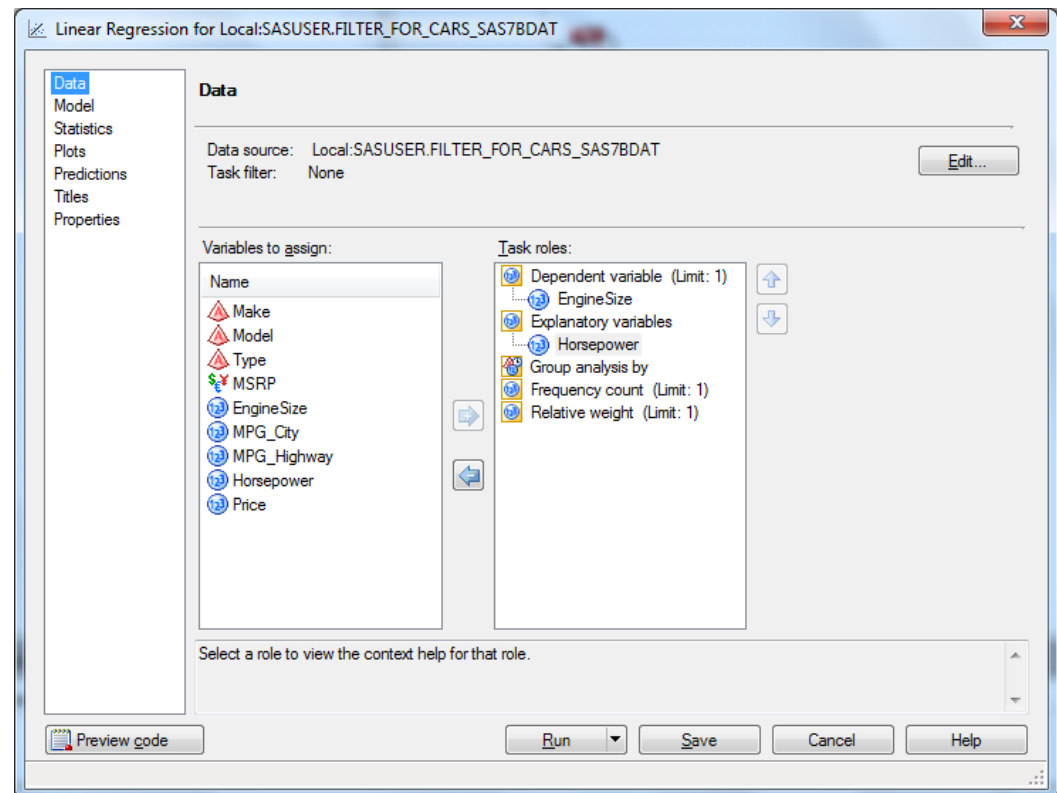
Teljesítmény

- Y: független változó

Motorméret

➤ Regresszió vizsgálat

- Task menü / Regression/ Linear reg...
- Data menüpont
 - Független változó:
Motor méret
 - Függő változó:
Teljesítmény
- Model menüpont
 - Maradjon változatlanul



> Regresszió eredménye

- Intercept – Y tengelymetszet
 - $B_0 = 0,616$ azaz ha $X=0$ azaz, ha a teljesítmény nulla, akkor a motor mérete 0,616 L azaz 600 cm³ (A modellben ezt nem értelmezzük)
- Meredekség
 - $B_1 = 0,01$ azaz ha a teljesítmény 1 kw nő, akkor a motor mérete 0,01 L (10 cm³)-rel nő
- Regressziós függvény
 - $\bar{y} = 0,616 + 0,01 * x$

Linear Regression Results

The REG Procedure

Model: Linear_Regression_Model

Dependent Variable: EngineSize Engine Size (L)

Number of Observations Read	123
Number of Observations Used	123

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	84.06564	84.06564	215.65	<.0001
Error	121	47.16915	0.38983		
Corrected Total	122	131.23480			

Root MSE	0.62436	R-Square	0.6406
Dependent Mean	3.20650	Adj R-Sq	0.6376
Coeff Var	19.47173		

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > t
Intercept	Intercept	1	0.61671	0.18512	3.33	0.0011
Horsepower		1	0.01028	0.00070012	14.68	<.0001

Köszönöm a figyelmet!