How about carsharing ?

ENAC second year econometrics project

Aurore Labrière Romain Fagoaga Patxi Artola

IENAC13TE

May, 2015

Outline

Outline

Introduction
Subject explanation

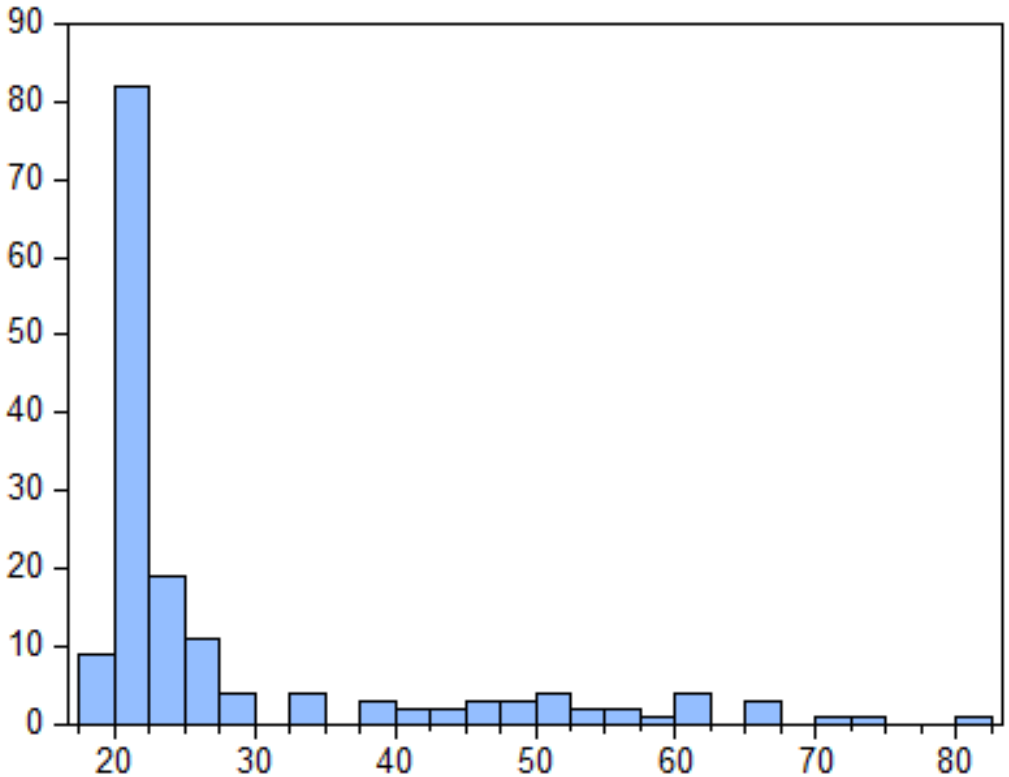
- People travel more and more using various means of transport
- Let us modelize why people like carsharing given their main features
- 161 answers to the Google survey with 30 questions
  Incentive: email feedback with the results in the report

Outline

First observations

- Answers from mostly ENAC students
- 52% men and 48% women
- Age mean: 28,6 years old and median: 22 years old
- Salary mean: 1400 euros and median: 1000 euros

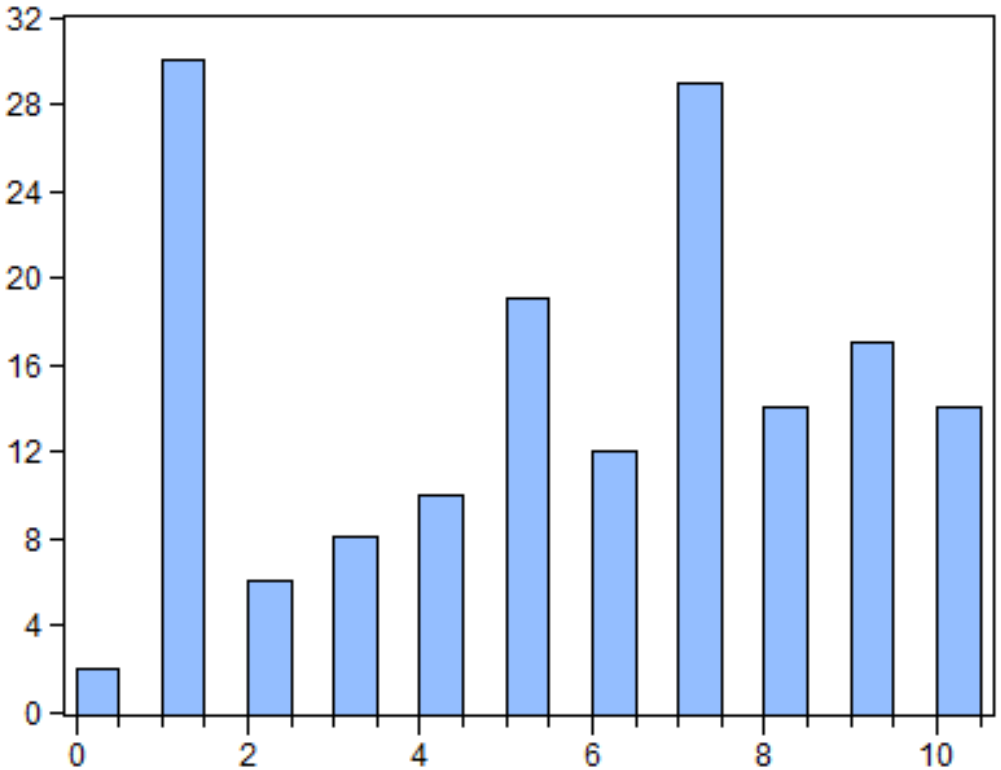
Age



Series: AGE	
Sample 1 161	
Observations 161	
Mean	28.63354
Median	22.00000
Maximum	82.00000
Minimum	18.00000
Std. Dev.	13.60684
Skewness	1.873454
Kurtosis	5.516676
Jarque-Bera	136.6688
Probability	0.000000

Fig.1 Age histogram and stats

Carsharing grades



Series: NOTE_COVOIT	
Sample 1 161	
Observations 161	
Mean	5.472050
Median	6.000000
Maximum	10.00000
Minimum	0.000000
Std. Dev.	3.027090
Skewness	-0.234882
Kurtosis	1.817085
Jarque-Bera	10.86727
Probability	0.004367

Fig.2 Carsharing grades histogram and stats

Age and carsharing

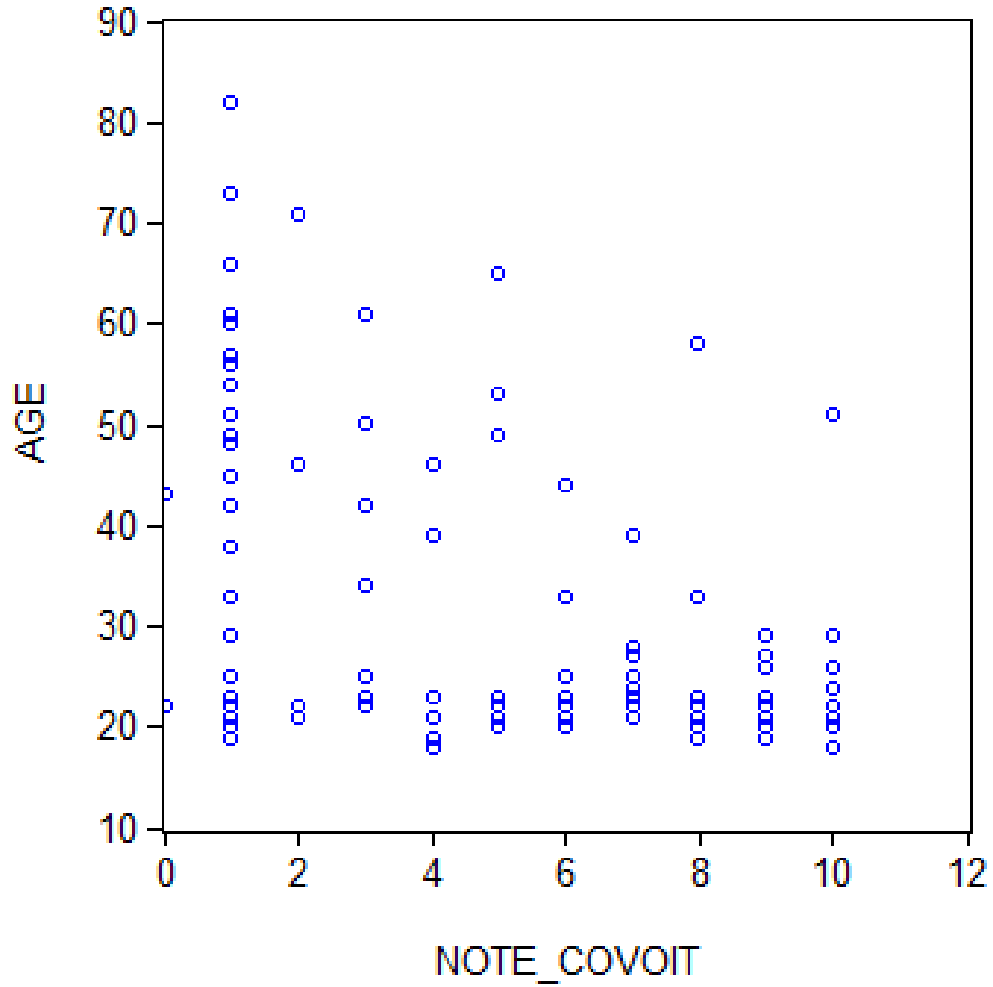


Fig.3 Age and carsharing

Outline

Refine variables

- Reverse causality of carpooling subscription
- Prevent perfect multicolinearity
- Correlation between variables

	BUDGET_MOY	AGE	RETRAITE	SALAIRE	EMPLOYE	ETUDIANT	ABO_TRAIN	NOTE_TRAIN
BUDGET_MOY	1,00	0,37	0,30	0,43	0,17	-0,31	-0,01	0,04
AGE	0,37	1,00	0,72	0,61	0,37	-0,75	-0,25	-0,22
RETRAITE	0,30	0,72	1,00	0,26	-0,14	-0,37	-0,12	-0,19
SALAIRE	0,43	0,61	0,26	1,00	0,60	-0,70	-0,21	-0,13
EMPLOYE	0,17	0,37	-0,14	0,60	1,00	-0,76	-0,25	-0,13
ETUDIANT	-0,31	-0,75	-0,37	-0,70	-0,76	1,00	0,26	0,19
ABO_TRAIN	-0,01	-0,25	-0,12	-0,21	-0,25	0,26	1,00	0,50
NOTE_TRAIN	0,04	-0,22	-0,19	-0,13	-0,13	0,19	0,50	1,00

Tab.1 Correlation matrix

Outline

First model

Equation 1

$$\widehat{NOTE\_COVOIT} = \widehat{\beta_0} + \widehat{\beta_1} AGE + \widehat{\beta_2} GENDER + \widehat{\beta_3} COUPLE +$$

$$\widehat{\beta_4} AUTOROUTE\_PROXI + \widehat{\beta_5} AMBIANCE + \widehat{\beta_6} CONFORT +$$

$$\widehat{\beta_7} DISPO + \widehat{\beta_8} FLEXIBILTE + \widehat{\beta_9} PRIX + \widehat{\beta_{10}} RAPIDITE +$$

$$\widehat{\beta_{11}} DISTANCE + \widehat{\beta_{12}} FREQUENCE + \widehat{\beta_{13}} VOITURE +$$

$$\widehat{\beta_{14}} POIDS\_BAGAGE + \widehat{\beta_{15}} ABO\_TRAIN +$$

$$\widehat{\beta_{16}} AVANTAGES\_PROMO$$

First model

Dependent Variable: NOTE_COVOIT
Method: Least Squares
Date: 05/19/15 Time: 21:56
Sample: 1 161
Included observations: 157

	Coefficient	Std. Error	t-Statistic	Prob.
C	4.941599	1.225656	4.031799	0.0001
AGE	-0.085208	0.018073	-4.714721	0.0000
HOMME	-0.065200	0.381412	-0.170945	0.8645
COUPLE	0.069124	0.408008	0.169419	0.8657
AUTOROUTE_PROXI	-0.321443	0.480422	-0.669086	0.5046
COVOIT_AMB	2.057205	0.411824	4.995352	0.0000
COVOIT_CONF	2.831528	0.897710	3.154166	0.0020
COVOIT_DISP	1.493417	0.467141	3.196932	0.0017
COVOIT_FLEX	1.888713	0.417538	4.523447	0.0000
COVOIT_PONCTU	0.040790	0.926751	0.044014	0.9650
COVOIT_PRIX	2.025896	0.623591	3.248757	0.0015
COVOIT_RAP	2.697229	0.743270	3.628866	0.0004
DIST_MOY	-0.000927	0.001100	-0.843154	0.4006
FREQUENCE	-0.015315	0.024904	-0.614970	0.5396
VOITURE	-0.111836	0.439019	-0.254741	0.7993
POIDS_MOY	-0.007849	0.016710	-0.469730	0.6393
ABO_TRAIN	-0.745727	0.391274	-1.905894	0.0587
AVANTAGES_PROMO	-0.043187	0.013287	-3.250367	0.0014

R-squared	0.511790	Mean dependent var	5.522293
Adjusted R-squared	0.452080	S.D. dependent var	2.996977
S.E. of regression	2.218410	Akaike info criterion	4.538986
Sum squared resid	684.0669	Schwarz criterion	4.889384
Log likelihood	-338.3104	Hannan-Quinn criter.	4.681295
F-statistic	8.571368	Durbin-Watson stat	1.957662
Prob(F-statistic)	0.000000		

Tab.2 First model for carsharing grade estimation

Effects of variables

Variables	Expected	Model
AGE	-	-
GENDER (H)	=	-
COUPLE	-	-
AUTOROUTE PROXI	+	-
AMBIANCE	+	+
CONFORT	-	+
DISPO	+	+
FLEXIBILITE	+	+
PONTUALITE	+	+
PRIX	+	+
RAPIDITE	-	+
DISTANCE MOY	-	=
FREQUENCE MOY	-	=
VOITURE	-	-
POIDS BAGAGES	-	=
ABO TRAIN	-	-
AVANTAGES PROMO	-	-

Second model

Equation 2

$$\widehat{NOTE\_COVOIT} = \widehat{\beta_0} + \widehat{\beta_1} AGE + \widehat{\beta_2} AMBIANCE + \widehat{\beta_3} CONFORT$$

$$+\widehat{\beta_4} DISPO + \widehat{\beta_5} FLEXIBILTE + \widehat{\beta_6} PRIX + \widehat{\beta_7} RAPIDITE$$

$$+\widehat{\beta_8} ABO\_TRAIN + \widehat{\beta_9} AVANTAGES\_PROMO$$

Second model

Dependent Variable: NOTE_COVOIT
Method: Least Squares
Date: 05/19/15 Time: 21:22
Sample: 1 161
Included observations: 161

	Coefficient	Std. Error	t-Statistic	Prob.
C	3.649337	0.857181	4.257370	0.0000
AGE	-0.078921	0.013615	-5.796443	0.0000
COVOIT_AMB	2.096678	0.378706	5.536429	0.0000
COVOIT_CONF	2.799382	0.856702	3.267626	0.0013
COVOIT_DISP	1.520805	0.433681	3.506741	0.0006
COVOIT_FLEX	1.892166	0.402959	4.695679	0.0000
COVOIT_PRIX	2.197307	0.575595	3.817456	0.0002
COVOIT_RAP	2.706745	0.687133	3.939187	0.0001
ABO_TRAIN	-0.739598	0.359134	-2.059394	0.0412
AVANTAGES_PROMO	-0.040418	0.012051	-3.353985	0.0010

R-squared	0.518885	Mean dependent var	5.472050
Adjusted R-squared	0.490209	S.D. dependent var	3.027090
S.E. of regression	2.161332	Akaike info criterion	4.439425
Sum squared resid	705.3745	Schwarz criterion	4.630817
Log likelihood	-347.3737	Hannan-Quinn criter.	4.517138
F-statistic	18.09491	Durbin-Watson stat	1.890855
Prob(F-statistic)	0.000000		

Tab.3 Second model for carsharing grade estimation

Second model

Equation 2

$$\widehat{NOTE\_COVOIT} = 3.649 - 0.079 * AGE + 2.097 * AMBIANCE$$

$$+2.799*CONFORT+1.521*DISPO+1.892*FLEXIBILTE$$

$$+2.197*PRIX+2.707*RAPIDITE-0.740*ABO\_TRAIN-$$

$$0.041*AVANTAGES\_PROMO$$

- Except age, personal features are not significant
- Transport preferences prevail
- Other transport advantages impact negatively

Outline

Residual from the model

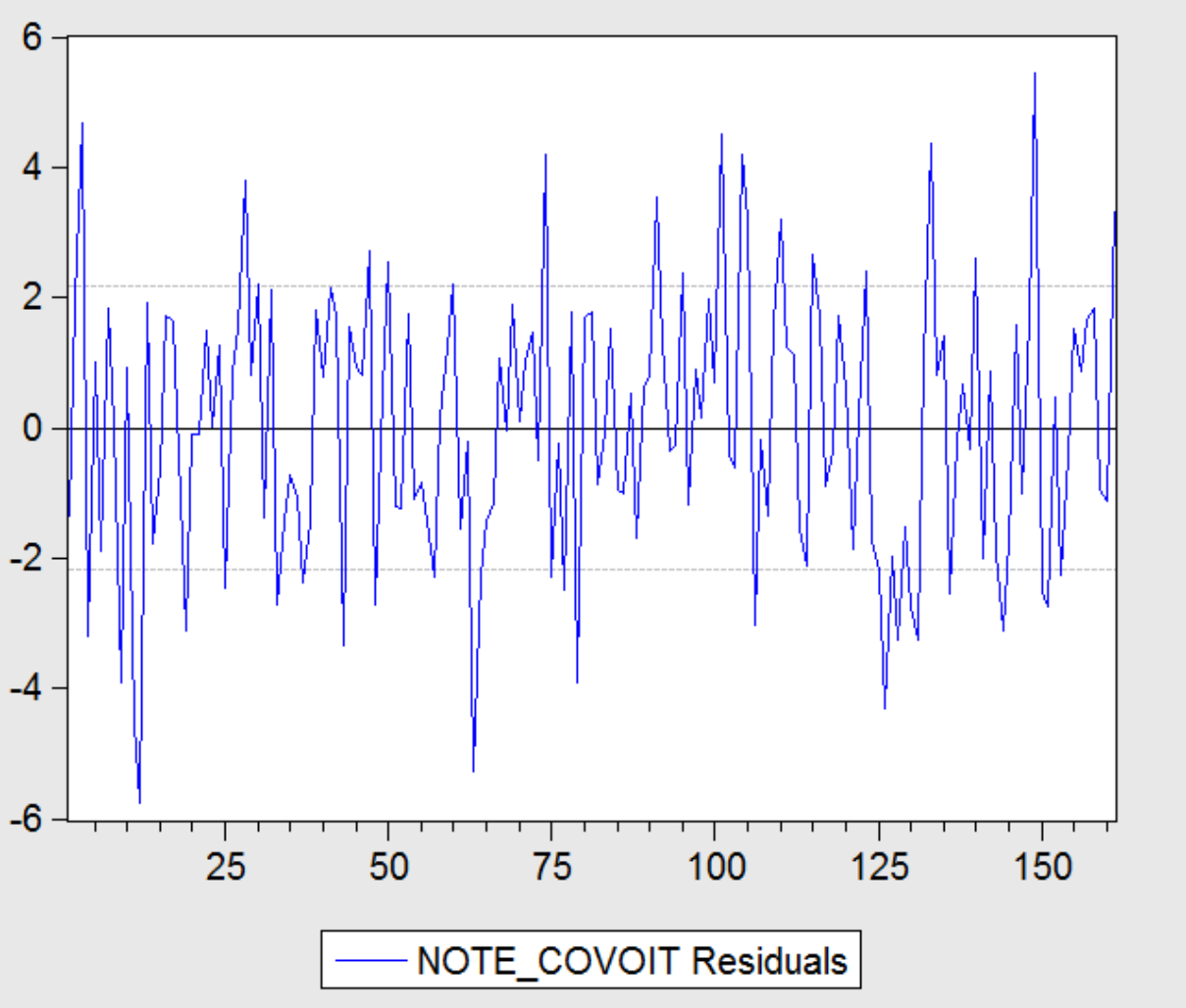


Fig.4 Residual graph

Homoscedasticity

The White test allows us to reject the heteroscedasticity hypothesis as the probability is greater than 0.05: it is **homoscedasctic**.

Heteroskedasticity Test: White

F-statistic	1.046929	Prob. F(40,120)	0.4126
Obs*R-squared	41.65024	Prob. Chi-Square(40)	0.3988
Scaled explained SS	32.75902	Prob. Chi-Square(40)	0.7848

Tab.4 Heteroscedasticity Test White

Residual normality

The Jarque-Bera test does not allow us to reject the **normality** hypothesis.

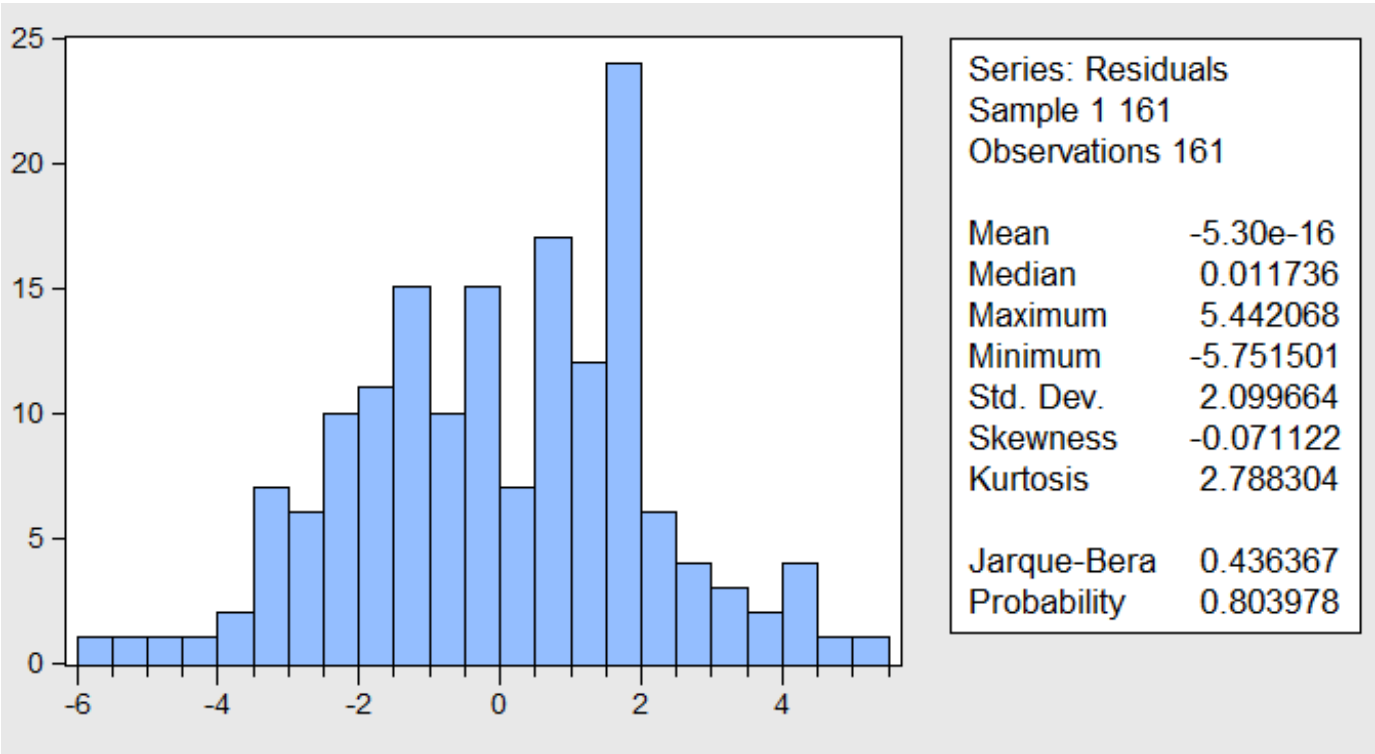


Fig.5 Normality test

Nonlinearity

The Ramsey RESET test **does not** allow us to **reject the null**.

Ramsey RESET Test:

F-statistic	1.198193	Prob. F(1,150)	0.2754
Log likelihood ratio	1.280951	Prob. Chi-Square(1)	0.2577

Tab.5 Ramsey RESET test

Outline

Plane model

Concerning the plane case, the same method gives us the following regression:

Sample: 1 161
Included observations: 159

	Coefficient	Std. Error	t-Statistic	Prob.
C	5.314211	0.457057	11.62701	0.0000
ABO_TRAIN	0.580332	0.349756	1.659252	0.0992
AVION_AMB	1.876930	0.746505	2.514290	0.0130
AVION_CONF	1.356839	0.369179	3.675292	0.0003
AVION_DISP	1.093537	0.617229	1.771689	0.0785
AVION_PRIX	1.864838	0.448772	4.155427	0.0001
AVION_INT	1.677565	0.403165	4.160995	0.0001
AVION_PONCTU	0.917860	0.366257	2.506055	0.0133
BUDGET__MOY	0.004041	0.002184	1.850068	0.0663
POIDS__MOY	-0.021697	0.011129	-1.949552	0.0531

R-squared	0.261910	Mean dependent var	7.238994
Adjusted R-squared	0.217327	S.D. dependent var	2.361235
S.E. of regression	2.088957	Akaike info criterion	4.372035
Sum squared resid	650.1973	Schwarz criterion	4.565048
Log likelihood	-337.5768	Hannan-Quinn criter.	4.450415
F-statistic	5.874699	Durbin-Watson stat	1.984296
Prob(F-statistic)	0.000001		

Tab.7 Plane grade estimation

The significant variables are different than the previous model.

Conclusion

Study conclusion

Our model gives a rather good estimation of the grade attributed to carpooling depending on age, preferences and particular advantages.

Opening

The study could be improved with more observations and also with more precise questions about carpooling.

Thank you for your attention. Have you got any questions ?