How important is TV in your life?

Laura GOUBERT- Elsa Paucsik- Xavier VanDen Bossche

ENAC

May 22nd, 2015

Introduction

Introduction

1. Aim : Highlighting the factors that influence the time you spend watching TV
2. How ? By spreading a survey to collect raw data

The survey

Method to collect data

- Creation of a survey on google drive
- spreading it on facebook and by e-mail

Results

206 answers in one week

Who answered ?

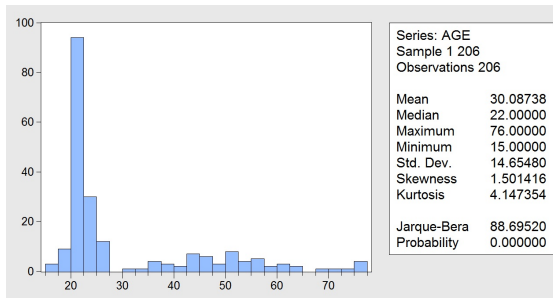1. Majority of people in their twenties
2. 57% of student

Figure: age repartition

Selection of the parameters

1. Beginning : 32 parameters
2. Selection of 26 parameters for the first model

several types of variable

- Majority of binary variables (Gendre, couple, owning a TV...)
- some discrete variables (activities' frequency..)

First model linear regression

Dependent Variable: TVTIME
Method: Least Squares
Date: 05/20/15   Time: 20:07
Sample: 1 206
Included observations: 206

	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.798792	5.372840	-0.148672	0.8820
ACTIVITIES	0.086853	0.259513	0.334677	0.7383
AGE	0.062874	0.079082	0.795055	0.4276
CARTOON	-0.416553	1.467051	-0.283939	0.7768
CHILD	0.314485	0.906441	0.346945	0.7290
CINEMA	-0.834188	1.830861	-0.455626	0.6492
CITY	0.573775	1.296376	0.442599	0.6586
DAY	0.580280	1.002120	0.579053	0.5633
EVENING	2.912740	1.947089	1.495946	0.1364
GENDER	0.168660	1.079645	0.156218	0.8760
INCOME	-0.000379	0.002669	-0.142018	0.8872
JOBFREE	-1.672111	4.171573	-0.400835	0.6890
MEAL	0.371081	1.039312	0.357045	0.7215
MOVIES	-0.110994	1.193168	-0.093025	0.9260
MOVIESTREAMING	-0.853792	1.275065	-0.669607	0.5040
NEWS	-0.095591	1.022727	-0.093467	0.9256
OBS	-0.003231	0.009830	-0.328718	0.7428
REALITYSHOW	1.386301	1.304474	1.062728	0.2893
SERIES	1.970603	1.060256	1.858610	0.0647
SINGLE	-1.515123	1.061430	-1.427436	0.1552
STUDENT	-3.937480	3.509368	-1.121991	0.2634
TELEVISION	3.533345	1.516704	2.329620	0.0209
TVGAMES	2.812070	1.312743	2.142134	0.0335
WEEK	2.181577	1.539535	1.417036	0.1582
WEEKEND	3.529347	1.574878	2.241028	0.0262
WORKER	-2.747340	3.550518	-0.773786	0.4401

R-squared	0.339296	Mean dependent var		9.524272
Adjusted R-squared	0.247532	S.D. dependent var		7.647534
S.E. of regression	6.633848	Akaike info criterion		6.739755
Sum squared resid	7921.428	Schwarz criterion		7.159778
Log likelihood	-668.1948	Hannan-Quinn criter.		6.909626
F-statistic	3.697470	Durbin-Watson stat		2.306248
Prob(F-statistic)	0.000000			

enhancement of the model

1. Correlation : Some high correlated variables
   - Age-Student (-0.65) /Income (-0.66)/Child(0.82)
   - Income-Student(-0.65)/Child(0.61)/Worker (0.6)

2. Many insignificant parameters

Final parameters

Finally we kept 8 parameters

Final model linear regression

Dependent Variable: TVTIME
Method: Least Squares
Date: 05/20/15   Time: 20:41
Sample: 1 206
Included observations: 206

	Coefficient	Std. Error	t-Statistic	Prob.
C	-4.986329	2.408431	-2.070364	0.0397
AGE	0.111287	0.035530	3.132221	0.0020
EVENING	2.981012	1.788647	1.666630	0.0972
REALITYSHOW	1.454521	1.194736	1.217441	0.2249
SERIES	1.825759	0.930946	1.961187	0.0513
SINGLE	-1.820945	0.965609	-1.885799	0.0608
TELEVISION	4.341005	1.331028	3.261393	0.0013
TVGAMES	2.799921	1.142437	2.450832	0.0151
WEEKEND	4.307504	1.462666	2.944967	0.0036

R-squared	0.315637	Mean dependent var	9.524272
Adjusted R-squared	0.287846	S.D. dependent var	7.647534
S.E. of regression	6.453695	Akaike info criterion	6.609889
Sum squared resid	8205.086	Schwarz criterion	6.755282
Log likelihood	-671.8186	Hannan-Quinn criter.	6.668691
F-statistic	11.35737	Durbin-Watson stat	2.320363
Prob(F-statistic)	0.000000		

1. $R^2 = 0.29$
2. Fstat$=0.000$

parameters	expected	actual
age	+	=
evening	+	+
realityshow	+	+
series	+	+
TVgames	+	+
single	-	-
television	+	+
weekend	+	+

Heteroskedasticity

Heteroskedasticity Test: White

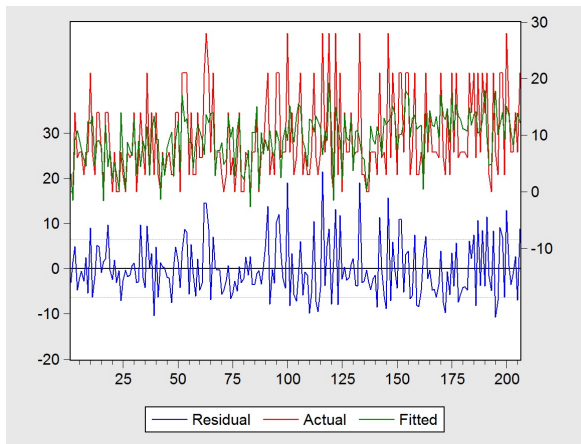|			
F-statistic	1.503561	Prob. F(35,170)	0.0469
Obs*R-squared	48.69485	Prob. Chi-Square(35)	0.0619
Scaled explained SS	53.02836	Prob. Chi-Square(35)	0.0259

Hypothesis of the null of homoskedasticity not rejected.

Normality test

	Series: Residuals
	Sample 1 206
	Observations 206
Mean	2.07e-16
Median	-1.382952
Maximum	21.46113
Minimum	-10.78960
Std. Dev.	6.326517
Skewness	0.866630
Kurtosis	3.381536
Jarque-Bera	27.03545
Probability	0.000001
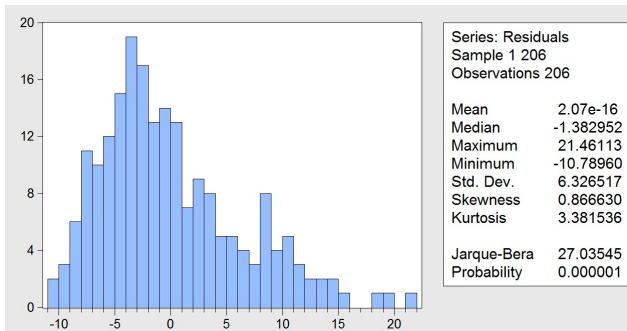
Hypothesis of normality of the error rejected.

Residual error

Model improvement

Limits of the model

- Too many binary variables
- Tv time not accurate enough
- Extrapolation of the values

Improvements

- Improve the accuracy of the question
- Ask more questions