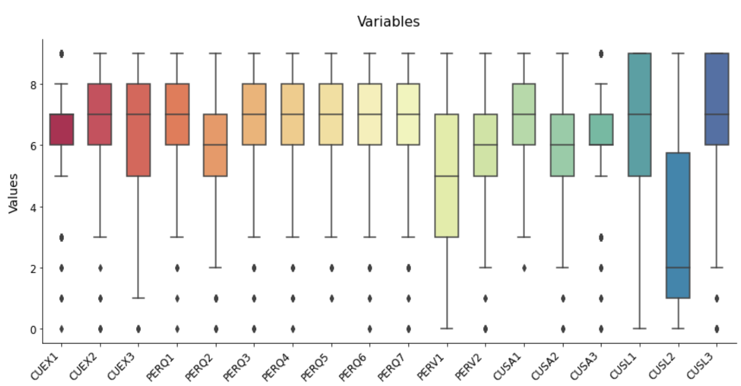
Modèles à équations structurelles : Théorie et applications

Travaux pratiques
By Manuel Esteban Arias

Question 1:

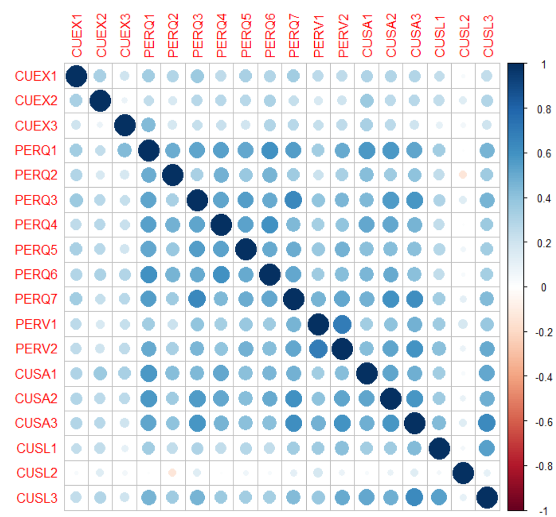
Explorer les données avec les statistiques descriptives et fournir un état des lieux de celles-ci.

The database is composed by eighteen variables rows that recorded the answer of 250 people to questions related to the expectations, perceived quality, perceived value, satisfaction and loyalty on the produces and services of a phone company. This questions, agreed in scales from 1 to 10 by the customers, present latent variables of the manifesto variables that permits us to estimate in which extend they may become frequent and loyal costumer of the phone services company.
Given that the date in this case is numeric, integer and within the same rank in all the variables, one fist approach to the date is the distribution, for this, the following boxplot has been made.



On it, we can see how Expectations variables present answers mainly high in with a median of 7 of two of them and the non extreme quartiles of the CUEX1 accumulated between 7 and 6. This is logic in the extend in which if the expectation where not high, people may not be negating the company's services.
The Perceptions on quality present similar behaviours with a median of seven again except for the PERQ2, suggestion that while human capital works well in the company, they still need to do technological improvements. The perceived values present high dispersion in both questions, specially in PERV1.Arround this, its s worth to be said that the customer's trade-off price-utility will depend on their relative appreciation to money, what may be linked to their revenue. As such, these variables will be extremely dispersed if the customer have diverse backgrounds, as is the case.
The costumer satisfaction present high values even when a bit lower than the resto of the variables seen until now. And finally, the Loyalty presents big dispersion. All the variables represent non atypic data from 1 to 10 and the CSL2 has a median close to 2, what, of course, depends on the

way the question was formulated (this 2 does not represent “agreeing level 2”, but a 2% of price differences of competitors.
In summary, we se that clients are often happy with the products and services of the company, they arrive with high expectations and they appreciate the quality received and they are in general satisfied, the appreciation of the value, though, is very diverse and they loyalty, while may be expect to be high, does not covers the lowest quartile of the population, which is exposed to the offers of the competitors.
Examined the nature of the data, it is worth to see they interact among them. For the, un useful the calculation of the correlation, as it can be seen in the graph above. There is the highest the correlation more intense the colour (indicating a direct or inverse correlation with blue or red respectively), and the biggest the circle. For far, it is possible to see that the variable Costumer Satisfaction, all for the PerceivedValue and some of the PerceivedQuality variables are all very correlated with CustomeLoyalty3(between 0,4 and 0,6 approx..), but not with the rest of CUSL. Given already some clues of which factors can explain the loyalty: the satisfaction, and even perception on quality and value.

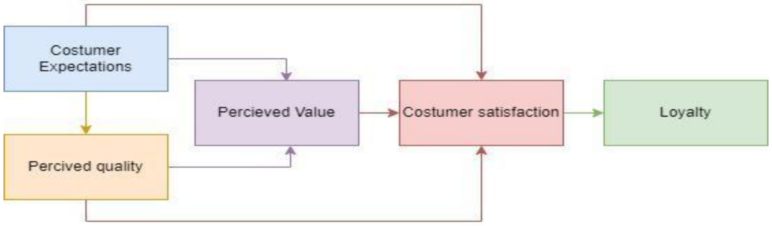


As well the Perceived Quality is internally correlated in most of their variables around 0,5 (not a surprise given the alignment on their distributions). This internal phenomenon is eve stronger with the perceived value, which presented an association of around 0,6. But, the most interesting in id the high correlation (around 0,6) of the customerSatisfaction with the quality version, for which shows high coefficient for most of the 7 variables.

Question 2 :

Le modèle établi par les experts est le modèle ECSI (figure 1). Cependant, dans le jeu de données fournit pour ce TP, les variables relatives aux plaintes et à l’image ont été éliminées. Le modèle est donc légèrement simplifié.
Spécifier ce modèle sous forme graphique et théorique avec l’ensemble des équations nécessaires.

The European Customer Satisfaction Index (ECSI) model presents correlations among manifesto variables that may be useful for forecasting the loyalty of clients. Given that our database omits some of them, the remaining external model would be like the following graph.



From the theoretical point of view the customer expectations can explain the perceived quality. Both, though, can explain the perceived values (in this regard, this is just a cost-benefit evaluation of the customer between the paid price and the non disappointing part of the quality received). This the variables will be able to explain the costumer satisfaction and, finally, the customer Satisfaction will explain the loyal.

Nevertheless, notions like expectations, perceptions, satisfaction, and loyalty are abstract subjective notions that cannot be measured precisely. Given this, within this model (the internal one) serval external models can be developed for proxying the real unobservable behaviors of these variables. As such, the internal models will be linear combinations of the data we got from the survey. All the parameters signaled will be estimated at once setting for a convergence in the variance of the target variable, the loyalty. As so the whole system of equations can be expressed as:

Intern Model

$$L = \gamma_1 CS + \zeta_1$$
$$CS = \beta_1 PV + \beta_2 CE + \beta_3\, PQ + \zeta_2$$
$$PV = \theta_1\, CE + \theta_2\, PQ + \zeta_3$$
$$PQ = \gamma_2\, CE + \zeta_3$$

Extern Model (formative type):

$$CE = \lambda_1 CUEX1 + \lambda_2 CUEX2 + \lambda_3 CUEX3 + \varepsilon_1$$
$$CS = \lambda_4 CUSA1 + \lambda_5 CUSA2 + \lambda_6 CUSA3 + \varepsilon_2$$
$$PQ = \lambda_7 PERQ1 + \lambda_8 PERQ2 + \lambda_9 PERQ3 + \lambda_{10} PERQ4 + \lambda_{11} PERQ5 + \lambda_{12} PERQ6 + \lambda_{13} PERQ7 + \varepsilon_3$$
$$PV = \lambda_{14} PERV1 + \lambda_{15} PERV2 + \varepsilon_5$$
$$L = \lambda_{16} CUSL1 + \lambda_{17} CUSL2 + \lambda_{18} CUSL3 + \varepsilon_6$$

In a theoretical way, and if data is scaled, the values of epsilon and zeta should be 0.

Question 3 :

Vérifier l'unidimensionnalité de chacun des blocs.

As mentioned before, the variables we aim to predict are not observable. In this regard the data we can measure will be latent variables serving for them. To run some of the models we are about to run, we will not collapse these variables into one and estimate the external model with these collapsed synthetic indicators. This process is commonly done though a PCA that leads to create dimensions that aim to explain the biggest part of the variance of the variables. Given that can only include one of the dimensions, we should make sure that the among of information reflects in this dimension significant, that explains most of the data. This implies verifying if the variance

explained of the first dimension, is significantly bigger than the one of the next one, or in other words, that the eigen value of the first dimension is significantly bigger than the other.
This will be verified by information blocks. These blocks are composed as can be seen in each one of the equations of the external model. One resulting bloc we can test:

```
> unidim_new(df,sat_blocks,alpha=0.05)
   Block MVs   C.alpha    DG.rho  eig.1st   eig.2nd    seuil      pval
1 block1   3 0.4350998 0.7250481 1.425152 0.9247435 1.175656 0.7994642
2 block2   7 0.8763201 0.9045427 4.033589 0.7850608 1.304245 0.9169201
3 block3   2 0.8174469 0.9163580 1.691256 0.3087440 1.124208 1.0000000
4 block4   3 0.7728313 0.8685839 2.063850 0.5256004 1.175656 0.9999999
5 block5   3 0.4801222 0.7329279 1.568738 0.9799160 1.175656 0.5886590
```

At first glance, we can see that the eigen values for the first dimension are always passing the unity, while the one of the second dimensions is never bigger than 1. From we can deduce directly a non despicable difference among the eigenvalues and therefore assume that the first dimension would explain most of the variance in the information block and as so be useful for including in a predictive model.

Question 4 :

Appliquer la méthode LISREL sur ce modèle à l'aide du package lavaan. Dérouler l'ensemble de la démarche de validation statistique vue en cours. Puis commenter de façon détaillée cette démarche et les résultats obtenus.

To apply the method LISREL we will run the system as specified in the Questions an apply a **G**eneralized **L**east **S**quared estimator given that all the variables have the same measurement unites. Said this, the input variable will look as follows.

```
# Modèle de mesure
PerQual =~ PERQ1 + PERQ2 + PERQ3 + PERQ4 + PERQ5 + PERQ6 + PERQ7
PerVal =~ PERV1 + PERV2
CusSat =~ CUSA1 + CUSA2 + CUSA3
Loyal =~  CUSL1 + CUSL2 + CUSL3
CusExp =~ CUEX1 + CUEX2 + CUEX3

# Modèle de structure
PerQual ~ CusExp
PerVal ~ CusExp + PerQual
CusSat ~ PerVal + CusExp + PerQual
Loyal ~ CusSat
```

The first model is not properly fitted according with the chi2 test, even when the rest of the test does not necessarily show a problem. The CFI and RMSEA show positive results for example. Let see if we can improve it.
For this, we can use the Modification index and include the first combination recommended it, as presented in the following screenshot:

```
> mod_ind_signif
      lhs op     rhs        mi        epc    sepc.lv   sepc.all   sepc.nox        p_val
1   PERQ1 ~~   CUEX3 20.159210  0.6730202  0.6730202  0.3186175  0.3186175 7.125632e-06
2   Loyal =~   CUSA3 15.111698  0.4750668  0.8655942  0.4361021  0.4361021 1.013334e-04
3  CusExp =~   CUSA1 14.564299  0.8194237  0.8126454  0.6101857  0.6101857 1.354568e-04
4   Loyal  ~ PerQual 13.678946 -1.1952582 -0.8001979 -0.8001979 -0.8001979 2.168723e-04
5  CusSat =~   PERQ7 12.714074  1.4063199  1.2871209  0.6169355  0.6169355 3.629140e-04
```

First, the variable CUSL2 was ommited (model 2)given the low corrilation detected in the eploratory analalysis. Latter, following this suggestions of the modification inde round by round it, was posible to add one by one contraints of covarians to the model. From model 4 , it was also addes a contratint to the covariance to the target vairables. .The iteration stoped whe addiing and additional contratin stopped changin sgnificatively the indexes. In this time this happened in the 3rd model (model 4) presented a worseing of the AIC. As so, the model 5 is the model 3, plus the constraint on the target variable.

Adequation Index	**Criteria**	**Model1**	**Model2**	**Model3**	**Model4**	**Model5**
Chi2	P-valeur > 0,05,	0.000	0.000	0.000	0.000	**0.000**
Comparative Fit Index (1989)	CFI > 0,9	0.928	0.932	0.942	0.937	**0.931**
Root Mean Square Error of Approximation	RMSEA ≤ 0,05.	0.066	0.068	0.063	0.066	**0.069**
Akaïke's Information Criterion (1974, 1987)	Min	17522.47	16230.41	16211.81	16221.04	**16231.57**
BIC (Bayesian Information Criterion)	Min	17673.89	16374.79	16359.71	16368.94	**16375.95**

As so, the final model, seen above, as the output of it.

```
# Modèle de mesure
PerQual =~ PERQ1 + PERQ2 + PERQ3+ PERQ4 + PERQ5 +
PERQ6 + PERQ7
PerVal =~ PERV1 + PERV2
CusSat =~ CUSA1 + CUSA2 + CUSA3
Loyal =~ CUSL1 + CUSL3
CusExp =~ CUEX1 + CUEX2 + CUEX3
# Modèle de structure
PerQual ~ CusExp
PerVal ~ CusExp + PerQual
CusSat ~ PerVal + CusExp + PerQual
Loyal ~ CusSat
# Covariance entre les erreurs
PERQ1 ~~ CUEX3
# Impose la variance de la fidélité à 0
Loyal ~~ 0*Loyal
```

Evaluations:

Phase 1: Here, several things can be said. Frist, even to the successive adjustments, t the model did not reached to demonstrate an elevated level of accuracy given that the Chi2 stayed always below 0,05 and as so, the null hypothesis of good adjustment is rejected. This inference is supported by the value RMSEA which passes 0,05 as well, suggesting a bad specification.

Phase 2: the latent variables table resulted to be al significant at the level of 5 % (suggesting that it was a good idea to include the constraints and to get rid of CUSL2). In this regard we have some realise by having evident that the latent variables can explains the manifesto one and that, as so the models for explain the manifesto variables have a good quality.

```
Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
  PerQual =~
    PERQ1             1.000                               1.197    0.768
    PERQ2             1.025    0.115    8.925    0.000    1.228    0.567
    PERQ3             1.290    0.106   12.208    0.000    1.545    0.750
    PERQ4             1.103    0.095   11.555    0.000    1.320    0.715
    PERQ5             0.931    0.084   11.047    0.000    1.114    0.687
    PERQ6             1.093    0.093   11.719    0.000    1.308    0.724
    PERQ7             1.298    0.107   12.109    0.000    1.553    0.745
  PerVal =~
    PERV1             1.000                               1.867    0.742
    PERV2             1.051    0.092   11.476    0.000    1.962    0.932
  CusSat =~
    CUSA1             1.000                               0.909    0.683
    CUSA2             1.682    0.158   10.649    0.000    1.530    0.751
    CUSA3             1.740    0.155   11.209    0.000    1.582    0.797
  Loyal =~
    CUSL1             1.000                               1.573    0.547
    CUSL3             1.112    0.136    8.190    0.000    1.750    0.717
  CusExp =~
    CUEX1             1.000                               1.034    0.561
    CUEX2             0.927    0.173    5.360    0.000    0.958    0.485
    CUEX3             0.848    0.191    4.434    0.000    0.876    0.374
```

```
Regressions:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
  PerQual ~
    CusExp            0.929    0.180    5.175    0.000    0.803    0.803
  PerVal ~
    CusExp           -0.022    0.335   -0.066    0.947   -0.012   -0.012
    PerQual           1.064    0.272    3.912    0.000    0.682    0.682
  CusSat ~
    PerVal            0.148    0.035    4.271    0.000    0.303    0.303
    CusExp            0.122    0.129    0.951    0.342    0.139    0.139
    PerQual           0.430    0.112    3.840    0.000    0.566    0.566
  Loyal ~
    CusSat            1.730    0.217    7.958    0.000    1.000    1.000

Covariances:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
 .PERQ1 ~~
   .CUEX3             0.680    0.160    4.257    0.000    0.680    0.313
```

Phase 3: Now, when the structural model is evaluated, we can see that the relationships suggested by the ECSI model are not describing in detail the relationships in the survey context.

In PerVal ~ CusExp + PerQual, CusExp seem to be a bad predictor for PerVal as well as in CusSat ~ PerVal + CusExp + PerQual. In contrast, the link between the perceptions is confirmed with the second higher coefficient in the model (1,064). The strongest relationship, though, is the one between the customer satisfaction and the loyalty, a positive new in the extend in which this variable is the one we want to produce. Nevertheless, seems to be a big part of the satisfaction that is not explains by the Preval, CusExp and PerQual.

Phase 4: When the error variances are analysis, there I only one problem in the P-value for the variable PERV1, whish passes 0,5, what suggests that the estimation of these variance is not significative.

When the manifesto variables are analysed, it can be se that while the variance of CusSat and PerQual not explains in the model is low, those of PerVal and CusExp are not. This, confirming the intuition that these two variables, related to the personal income of the costumers, present a higher dispersion when the customers come from different background.

```
Variances:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
   .Loyal             0.000                               0.000    0.000
   .PERQ1             1.000    0.105    9.533    0.000    1.000    0.411
   .PERQ2             3.180    0.299   10.642    0.000    3.180    0.678
   .PERQ3             1.858    0.191    9.709    0.000    1.858    0.438
   .PERQ4             1.669    0.167    9.986    0.000    1.669    0.489
   .PERQ5             1.388    0.137   10.160    0.000    1.388    0.528
   .PERQ6             1.557    0.157    9.923    0.000    1.557    0.476
   .PERQ7             1.940    0.199    9.756    0.000    1.940    0.446
   .PERV1             2.852    0.341    8.364    0.000    2.852    0.450
   .PERV2             0.582    0.255    2.281    0.023    0.582    0.131
   .CUSA1             0.947    0.094   10.037    0.000    0.947    0.534
   .CUSA2             1.805    0.191    9.462    0.000    1.805    0.435
   .CUSA3             1.436    0.162    8.838    0.000    1.436    0.364
   .CUSL1             5.801    0.546   10.629    0.000    5.801    0.701
   .CUSL3             2.892    0.295    9.788    0.000    2.892    0.486
   .CUEX1             2.324    0.277    8.389    0.000    2.324    0.685
   .CUEX2             2.983    0.315    9.477    0.000    2.983    0.765
   .CUEX3             4.733    0.457   10.364    0.000    4.733    0.860
   .PerQual           0.510    0.167    3.059    0.002    0.356    0.356
   .PerVal            1.909    0.316    6.037    0.000    0.547    0.547
   .CusSat            0.136    0.037    3.647    0.000    0.165    0.165
    CusExp            1.069    0.287    3.724    0.000    1.000    1.000
```

```
$summary
                          cov
srmr                    0.050
srmr.se                 0.003
srmr.exactfit.z         6.422
srmr.exactfit.pvalue    0.000
usrmr                   0.038
usrmr.se                0.005
usrmr.ci.lower          0.030
usrmr.ci.upper          0.046
usrmr.closefit.h0.value 0.050
usrmr.closefit.z       -2.456
usrmr.closefit.pvalue   0.993
```

Phase 5: The verification of the residuals can be made evaluation their covariances across the variables a though different test. Here we test the significance for the Standardized Root Mean Square Residuals and to the. As we can see, the SRMR is lower than 0,8 an a so, the model has a good adjustment. Similarly, the Unbiased SRMR does not passe this threshold and as so, at least from the residuals point of view, we can say that the model has a correct specification.

Question 5:

Si l'on désire appliquer sur ce modèle, l'approche PLS : quelles sont les différences majeures entre cette approche et la méthode LISREL, en termes de spécification du modèle et de paramètres à estimer.

While seeking the same, these two different algorithms differ in method and objectives. By one side, the biggest difference may be that the LISREL approach seeks to estimate ML estimators, while the PLS keeps the OLS modelling. Nevertheless, the understanding of the optimization algorithm may be more revenant. Here we will have a model of the external model in which the variance will be capture by a "error" variable Zeta that will have its own estimation that will be calculated as well.

The absence of this extra parameter in the PLS permits the convergence not to be reached by the constraining the covariance matrix of the residuals as in LISREL, but by the minimizing the trace of the matrix of error's variance of the manifesto variables. As so, the equations that constrains the residuals are no longer necessaries and the estimators calculated decrease. The PSL converges, in that regard, can be consider as natural while in LISREL will depend on don the constraints we identify, is instable.

Question 6:

Utiliser le package plspm pour ajuster ce modèle en choisissant. Dérouler l'ensemble de la démarche de validation statistique vue en cours et commenter de façon détaillée les résultats obtenus.

```
> summary(satpls3)
PARTIAL LEAST SQUARES PATH MODELING (PLS-PM)

----------------------------------------------------------
MODEL SPECIFICATION
1   Number of Cases      250
2   Latent Variables     5
3   Manifest Variables   17
4   Scale of Data        Standardized Data
5   Non-Metric PLS       FALSE
6   Weighting Scheme     centroid
7   Tolerance Crit       1e-06
8   Max Num Iters        100
9   Convergence Iters    3
10  Bootstrapping        TRUE
11  Bootstrap samples    100

----------------------------------------------------------
BLOCKS DEFINITION
      Block         Type   Size   Mode
1   CusExpl    Exogenous      3      A
2    PerQua   Endogenous      7      A
3    PerVal   Endogenous      2      A
4    CusSat   Endogenous      3      A
5     Loyal   Endogenous      2      A

----------------------------------------------------------
BLOCKS UNIDIMENSIONALITY
         Mode  MVs  C.alpha  DG.rho  eig.1st  eig.2nd
CusExpl     A    3    0.435   0.725     1.43    0.925
PerQua      A    7    0.876   0.905     4.03    0.785
PerVal      A    2    0.817   0.916     1.69    0.309
CusSat      A    3    0.773   0.869     2.06    0.526
Loyal       A    2    0.707   0.872     1.55    0.454
```

To test the model already specified by used the PLS approach a model type A with a weight scheme "centroid" was estimated. While one of the biggest advantages of this methods (and the reason why is the most used ones) is that it is permitted to process missing data, however, since our database is complete, this is an irrelevant featured.

Said this the algorithm starts by estimating the eternal models' loadings for the difference information blocks, later, process the internal model with a centroid scheme, it estimates the eternal models given the results for the interna, later again the external ones, and so on until the model struct converge to stable estimators. Only at the end, each one of the variables manifestos will be regress on their explicative latent variables to assess their capacity to predict them.

Once again, the **dimensionality** of the information blocks can be verified by realising that all the first eigen values are higher than one, all the second or them have a value below the unity.

OUTER MODEL

		weight	loading	communality	redundancy
CusExpl					
1	CUEX1	0.529	0.774	0.600	0.000
1	CUEX2	0.478	0.689	0.474	0.000
1	CUEX3	0.443	0.589	0.347	0.000
PerQua					
2	PERQ1	0.215	0.808	0.653	0.191
2	PERQ2	0.146	0.630	0.397	0.116
2	PERQ3	0.204	0.783	0.613	0.179
2	PERQ4	0.177	0.770	0.593	0.173
2	PERQ5	0.177	0.747	0.558	0.163
2	PERQ6	0.179	0.786	0.617	0.180
2	PERQ7	0.214	0.773	0.598	0.175
Perval					
3	PERV1	0.490	0.902	0.814	0.281
3	PERV2	0.596	0.935	0.875	0.302
CusSat					
4	CUSA1	0.384	0.798	0.637	0.421
4	CUSA2	0.395	0.846	0.716	0.472
4	CUSA3	0.426	0.843	0.710	0.469
Loyal					
5	CUSL1	0.494	0.844	0.712	0.318
5	CUSL3	0.641	0.911	0.829	0.371

As expected, we obtain first the **loadings** estimates for each manifesto variable derived from its relationships with the latent variables. The outer model table shows communality ratios high enough for signalling a high-quality model with exception of CUEX2 and CUEX3, which index is the furthest from one. this may suggest that the outer model of Costumer expectations is has not a high quality. This is confirmed by the redundancy indices that, as zero, suggest that the information block is not providing of useful information properly the structural model. In the rest of the case, the variables respond well to the latent variables while not very pertinently to the inner structure.

Second, we obtain the **external weights** of the latent variables on its information blocks. We can see that most of the intercept are not significant (this given the data scaled performed with the algorithm). In contrast most of the variables are declared as significant in the structural models what give point to eh ECSI specification. The only exception is the on of the Perceive Quality in the Loyal equation. This is not only hard to explain but counter-intuitive in the extend in which as a costumer, the quality should be big reason for coming back to the product or service provider.

INNER MODEL

$PerQua

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	3.42e-17	0.0534	6.40e-16	1.00e+00
CusExpl	5.41e-01	0.0534	1.01e+01	2.24e-20

$Perval

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	-1.36e-16	0.0515	-2.64e-15	1.00e+00
CusExpl	6.91e-02	0.0612	1.13e+00	2.60e-01
PerQua	5.47e-01	0.0612	8.94e+00	9.19e-17

$CusSat

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	1.09e-16	0.0372	2.94e-15	1.00e+00
CusExpl	9.75e-02	0.0443	2.20e+00	2.87e-02
PerQua	5.95e-01	0.0508	1.17e+01	1.94e-25
Perval	2.34e-01	0.0459	5.09e+00	7.25e-07

$Loyal

	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	9.83e-17	0.0475	2.07e-15	1.00e+00
CusExpl	5.57e-02	0.0572	9.75e-01	3.31e-01
PerQua	-2.60e-02	0.0810	-3.20e-01	7.49e-01
Perval	1.88e-01	0.0617	3.05e+00	2.54e-03
CusSat	5.27e-01	0.0815	6.46e+00	5.44e-10

CROSSLOADINGS

		CusExpl	PerQua	Perval	CusSat	Loyal
CusExpl						
1	CUEX1	0.774	0.420	0.284	0.361	0.278
1	CUEX2	0.689	0.336	0.203	0.362	0.310
1	CUEX3	0.589	0.356	0.266	0.317	0.182
PerQua						
2	PERQ1	0.495	0.808	0.468	0.672	0.466
2	PERQ2	0.310	0.630	0.295	0.474	0.350
2	PERQ3	0.428	0.783	0.467	0.641	0.452
2	PERQ4	0.367	0.770	0.400	0.602	0.355
2	PERQ5	0.393	0.747	0.467	0.503	0.362
2	PERQ6	0.432	0.786	0.431	0.550	0.336
2	PERQ7	0.420	0.773	0.538	0.683	0.451
Perval						
3	PERV1	0.324	0.476	0.902	0.499	0.416
3	PERV2	0.346	0.589	0.935	0.625	0.528
CusSat						
4	CUSA1	0.480	0.623	0.422	0.798	0.500
4	CUSA2	0.402	0.677	0.514	0.846	0.490
4	CUSA3	0.379	0.651	0.591	0.843	0.622
Loyal						
5	CUSL1	0.293	0.387	0.422	0.461	0.844
5	CUSL3	0.361	0.524	0.484	0.660	0.911

The third point to verify is the **correlations of manifesto variables** with each of the latent variable's ones that. It is expected that this correlation is higher between the manifesto variables and the latent variable that englobes the information block. As we can see, this principle is respected, and the variables present higher coefficient int he columns of the latent variable associated to their information block. This is still a good sing about the model specification.

rsq

	original	Mean.Boot	Std.Error	perc.025	perc.975
PerQua	0.292	0.308	0.0573	0.218	0.424
Perval	0.345	0.352	0.0621	0.234	0.491
CusSat	0.660	0.673	0.0407	0.605	0.744
Loyal	0.447	0.464	0.0600	0.349	0.584

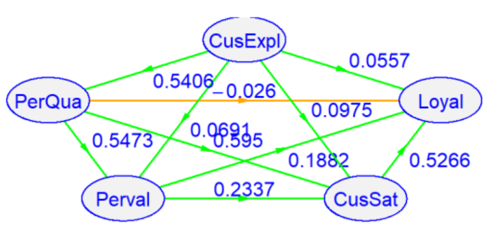
The **bootstrap validation** shows that the mean of the T value increase marginally what question the necessity of using bootstrapping but what does not represent ahigh loss for the research and therefore does not question the obtained model itself.

Going back to the inner model, the Costumer satisfactions highlights as the one who contributes the more to the model accuracy with a R2 contribution of 0,6 just followed by the Loyalty, the target variable. The communality holds high values as well suggesting the already known quality of mediation in the external variables, in contrast the redundancy, the one talks about the structure quality is low, seeding doubt about the overall specification. Nevertheless, the **Goodness of fit** is high enough (0,55) for thinking that the model has the minimum requires for a good fitting.

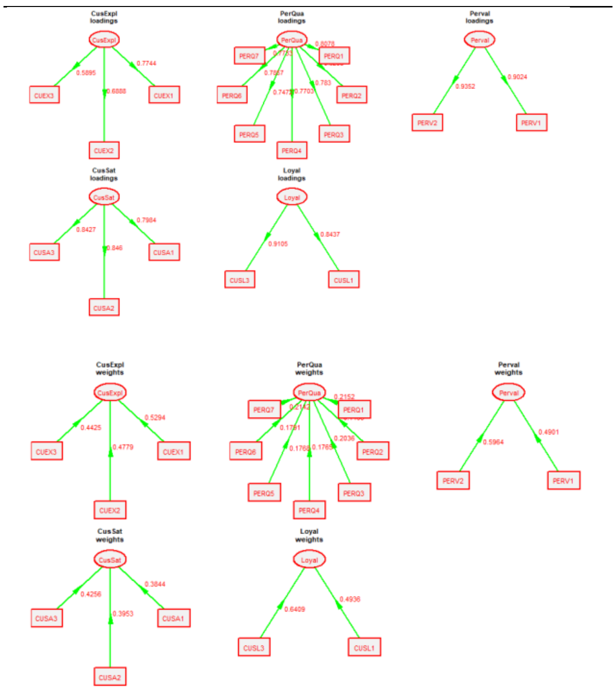
```
-------------------------------------------------------------
SUMMARY INNER MODEL
                Type     R2  Block_Communality  Mean_Redundancy    AVE
CusExpl    Exogenous  0.000              0.474            0.000  0.474
PerQua    Endogenous  0.292              0.576            0.168  0.576
Perval    Endogenous  0.345              0.844            0.291  0.844
CusSat    Endogenous  0.660              0.688            0.454  0.688
Loyal     Endogenous  0.447              0.770            0.345  0.770

-------------------------------------------------------------
GOODNESS-OF-FIT
[1]  0.5251
```

Inner Model



Outer Model



Question 7 :

Appliquer sur ce même modèle, l'approche RFPC.

The Regression on the First Principal Components (RFPC) takes advantage of the notion of dimensionality for run the internal model based on external model that defines on the first component of a PCA on the manifesto variables. These Linear transformations permit to have 1:1 external model with the risk components associated so the loadings are just projections of the eigen values of the components.

[1] "OUTER MODEL"

	name	block	weight	loading	communality	redundancy
CUEX1	CUEX1	CusExpl	-0.6687715	-0.7983776	0.6374068	0.0000000
CUEX2	CUEX2	CusExpl	-0.6020303	-0.7187021	0.5165328	0.0000000
CUEX3	CUEX3	CusExpl	-0.4362387	-0.5207806	0.2712124	0.0000000
PERQ1	PERQ1	PerQua	-0.3994123	-0.8021716	0.6434792	0.1825744
PERQ2	PERQ2	PerQua	-0.3190514	-0.6407763	0.4105942	0.1164979
PERQ3	PERQ3	PerQua	-0.3849215	-0.7730686	0.5976350	0.1695670
PERQ4	PERQ4	PerQua	-0.3896469	-0.7825589	0.6123985	0.1737559
PERQ5	PERQ5	PerQua	-0.3743113	-0.7517593	0.5651420	0.1603478
PERQ6	PERQ6	PerQua	-0.3960595	-0.7954378	0.6327214	0.1795221
PERQ7	PERQ7	PerQua	-0.3764501	-0.7560548	0.5716188	0.1621855
PERV1	PERV1	Perval	0.7071068	0.9195803	0.8456280	0.2821821
PERV2	PERV2	Perval	0.7071068	0.9195803	0.8456280	0.2821821
CUSA1	CUSA1	CusSat	0.5576325	0.8011009	0.6417626	0.4192268
CUSA2	CUSA2	CusSat	0.5939150	0.8532247	0.7279924	0.4755558
CUSA3	CUSA3	CusSat	0.5799232	0.8331239	0.6940955	0.4534129
CUSL1	CUSL1	Loyal	0.7071068	0.8792610	0.7730999	0.3324096
CUSL3	CUSL3	Loyal	0.7071068	0.8792610	0.7730999	0.3324096

The results found in the case are not very different from the ones of PSL in the extend in which the only component with internal problems is still Consumer expectation. While their contribution to the structure model remains almost zero, their measurement specification passes form the 50% often but remains low (Perceived value, one's gain, shows mediocre results for the redundancy). Once again, thought, for most of the assess the external equations seem to have a good capacity to measure but the introduction to the structural problem seem to represent a problem given hat the redundancy never passe from 0,47.

When the correlations between s the block and the manifesto variables are reviewed, the signs for CusEpxl and PerQual are now negative. Nevertheless, the magnitude keeps responding to the information block the one the variables belong and as so, the assumption on the correct specification of the information blocks remains correct.

[1] "CROSSLOADINGS"

	CusExpl	PerQua	Perval	CusSat	Loyal
CUEX1	-0.7983776	-0.4171494	0.2842972	0.3605958	0.2781806
CUEX2	-0.7187021	-0.3378043	0.2009819	0.3624666	0.3074586
CUEX3	-0.5207806	-0.3519767	0.2658184	0.3187816	0.1757253
PERQ1	-0.4778783	-0.8021716	0.4603541	0.6729638	0.4585318
PERQ2	-0.3099198	-0.6407763	0.2895661	0.4731244	0.3430153
PERQ3	-0.4282930	-0.7730686	0.4639693	0.6399254	0.4415384
PERQ4	-0.3644111	-0.7825589	0.3970493	0.6026270	0.3475585
PERQ5	-0.3950445	-0.7517593	0.4621458	0.5029647	0.3594370
PERQ6	-0.4269534	-0.7954378	0.4273948	0.5513254	0.3307172
PERQ7	-0.4149812	-0.7560548	0.5361198	0.6820993	0.4454432
PERV1	-0.3183733	-0.4715703	0.9195803	0.4972502	0.4152487
PERV2	-0.3395962	-0.5848688	0.9195803	0.6226063	0.5222596
CUSA1	-0.4739644	-0.6219552	0.4184891	0.8011009	0.4884475
CUSA2	-0.3970304	-0.6711740	0.5084349	0.8532247	0.4816922
CUSA3	-0.3794577	-0.6440419	0.5852766	0.8331239	0.6088844
CUSL1	-0.2992446	-0.3833261	0.4187473	0.4586503	0.8792610
CUSL3	-0.3603822	-0.5195652	0.4776556	0.6571748	0.8792610

[1] "INNER MODEL"

	Estimate	Std. Error	t value	Pr(>|t|)
PerQua	NA	NA	NA	NA
CusExpl	0.89612364	0.09041235	9.9115181	1.007059e-19
Perval	NA	NA	NA	NA
CusExpl	-0.07876111	0.06685343	-1.1781163	2.398841e-01
PerQua	-0.34701193	0.03973822	-8.7324468	3.823334e-16
CusSat	NA	NA	NA	NA
CusExpl	-0.12490614	0.05353430	-2.3331984	2.044442e-02
PerQua	-0.42184894	0.03630151	-11.6207008	3.578429e-25
Perval	0.25739493	0.05080929	5.0659028	7.982975e-07
Loyal	NA	NA	NA	NA
CusExpl	-0.07006102	0.06018656	-1.1640643	2.455303e-01
PerQua	0.01545508	0.05024088	0.3076196	7.586332e-01
Perval	0.18592325	0.05937527	3.1313248	1.951575e-03
CusSat	0.43437761	0.07090011	6.1266141	3.555012e-09

Now lets analyses the structural model. I the output the algorithm managed to fit all the variables. This time, the costume experience is inversely correlated to the loyalty rate as well as for the customer satisfaction and the Perceived value. While this affirmation may be not sensible, in the model may be present problems for the interpretation, however, its accuracy may still be remarkable.

Finally, the evaluation of the model hosts the R2, communality and redundancy columns that can be compared with the one of the PLS approach. Frist, the R2 is slightly smaller e in general suggesting similar contribution form he different blocks to the structural model,

```
[1] "SUMMARY INNER MODEL"
                R2 Mean communality Mean Redundancy
CusExpl 0.0000000         0.4750507       0.0000000
CusSat  0.2837301         0.6879501       0.4493985
Loyal   0.3336953         0.7730999       0.3324096
PerQua  0.6532429         0.5762270       0.1634929
Perval  0.4299697         0.8456280       0.2821821
[1] "GOODNESS-OF-FIT"
[1] 0.5343532
```

Nevertheless, the order in which the components contribute is now different, and this may be link to the change of signs we saw I the correlation matrix. Second, the Communality presents the high increase in all the components suggesting the working with reduced dimension permit to capture better the variance of the manifesto variables and link better the latent variables. Finally, the redundancy increases as well slightly. Finally, the goodness to fit reflex this improvement in the quality ratios by increasing a little bit the final accuracy score.

Question 8 :

Comparer tout d'abord les pouvoirs explicatifs des trois modèles de mesure et des trois modèles de structure (LISREL, PLS et RFPC). Que remarquez-vous ?

PLS

```
----------------------------------------------------------
SUMMARY INNER MODEL
              Type     R2  Block_Communality  Mean_Redundancy    AVE
CusExpl  Exogenous  0.000              0.474            0.000  0.474
PerQua  Endogenous  0.292              0.576            0.168  0.576
Perval  Endogenous  0.345              0.844            0.291  0.844
CusSat  Endogenous  0.660              0.688            0.454  0.688
Loyal   Endogenous  0.447              0.770            0.345  0.770

----------------------------------------------------------
GOODNESS-OF-FIT
[1]  0.5251
```

RFPC


```
[1] "SUMMARY INNER MODEL"
                R2 Mean communality Mean Redundancy
CusExpl 0.0000000         0.4750507       0.0000000
CusSat  0.2837301         0.6879501       0.4493985
Loyal   0.3336953         0.7730999       0.3324096
PerQua  0.6532429         0.5762270       0.1634929
Perval  0.4299697         0.8456280       0.2821821
[1] "GOODNESS-OF-FIT"
[1] 0.5343532
```


The PLS and the RFPC approaches are very similar in terms of effectivity. Both presents a high Goodness of fit. Nevertheless, the RFPC presents slightly better results not only in terms of fitting but int herm s of communality and redundancy. As so we can the that a dimension reduction with PCA methods can drive better the date and extract form it the most useful information for the model. In terms of predictability is therefore superior.

However, the data transformation applied this method makes more complicated the analysis o the information in the extend in which the manifesto variables are no longer observable, the latent variable appears and while useful for predict says nothing about the reality int herm so short run changes.

The implementation of analgorithm like this one has higher perfomance but implies to work with thanform data which may complicate the struchral change exlanationa snd as dso would make difficult to react in terms od daily life entrepreneurial activity.

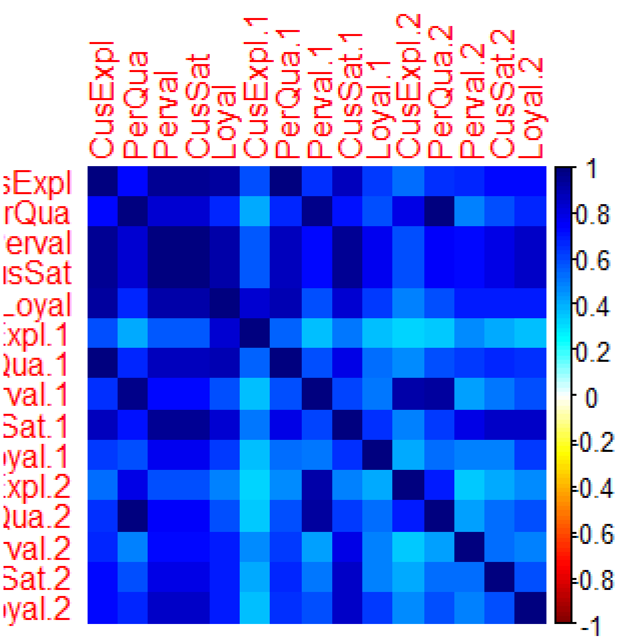
Question 9 :

Construire les variables latentes et les poids externes normalisés à l'aide de l'approche exploratoire pour LISREL, puis comparer les variables latentes et les poids externes normalisés issus des trois méthodes : LISREL, PLS et RFPC.

```
> outer_model_3_methodes
```

	lat_var	man_var	lambda_LISREL	deno_lambda	lambda_stand_LISREL	lambda	lambda_PLS	lambda_stand_PLS	lambda_RFPC	lambda_stand_RFPC
1	CusExp	CUEX1	1.0000000	1.605338	0.6229	1.0000	0.7744	0.6495	-0.7984	-0.6688
2	CusExp	CUEX2	0.9266013	1.605338	0.5772	0.9266	0.6888	0.5777	-0.7187	-0.6020
3	CusExp	CUEX3	0.8476563	1.605338	0.5280	0.8477	0.5895	0.4944	-0.5208	-0.4362
4	CusSat	CUSA1	1.0000000	2.618684	0.3819	1.0000	0.7984	0.5559	0.8011	0.5576
5	CusSat	CUSA2	1.6823176	2.618684	0.6424	1.6823	0.8460	0.5889	0.8532	0.5939
6	CusSat	CUSA3	1.7399174	2.618684	0.6644	1.7399	0.8427	0.5867	0.8331	0.5799
7	Loyal	CUSL1	1.0000000	1.495565	0.6686	1.0000	0.8437	0.6797	0.8793	0.7071
8	Loyal	CUSL3	1.1120761	1.495565	0.7436	1.1121	0.9105	0.7335	0.8793	0.7071
9	PerQual	PERQ1	1.0000000	2.945709	0.3395	1.0000	0.8078	0.4024	-0.8022	-0.3994
10	PerQual	PERQ2	1.0254811	2.945709	0.3481	1.0255	0.6298	0.3137	-0.6408	-0.3191
11	PerQual	PERQ3	1.2903207	2.945709	0.4380	1.2903	0.7830	0.3901	-0.7731	-0.3849
12	PerQual	PERQ4	1.1028367	2.945709	0.3744	1.1028	0.7703	0.3837	-0.7826	-0.3897
13	PerQual	PERQ5	0.9306939	2.945709	0.3159	0.9307	0.7472	0.3723	-0.7518	-0.3743
14	PerQual	PERQ6	1.0929513	2.945709	0.3710	1.0930	0.7857	0.3914	-0.7954	-0.3960
15	PerQual	PERQ7	1.2975660	2.945709	0.4405	1.2976	0.7733	0.3852	-0.7561	-0.3765
16	PerVal	PERV1	1.0000000	1.450563	0.6894	1.0000	0.9024	0.6944	0.9196	0.7071
17	PerVal	PERV2	1.0507777	1.450563	0.7244	1.0508	0.9352	0.7196	0.9196	0.7071

When the loading of all the estimations algorithms compered, the estimators of the manifest variables. Comparing the standardized lambdas, the LisRel holds positive values for all the coefficients, what goes along with the theory of consumption assuming that positive perceptions may rise the probability of coming back to purchase the company's products and services. Only the costumer experiences and satisfaction will remain positive in their coefficients. Other interesting fact is that the PLS and RFPC holds higher magnitudes in the estimators, probability for compensation the asymmetries in the signs.



Even when the data treated in the PLS and RFPC algorithms is supposed to be, in some extend, "filtered" for explaining better the model, it is interesting to se that the correlations within the LISREL structural model are stronger. Even when the adjustment of the model can be smaller the constraining of the covariance matrix actual approach the data in a hard but effective way. In contrast, the dimension reduction permits the covariance to be softer while getting element nor for reconstruction/explaining the variable but uniquely for predicting it.