The CONTENTS Procedure

Data Set Name	MYDATA.PANEL95_TABLE	**Observations**	10548
Member Type	DATA	**Variables**	6
Engine	V9	**Indexes**	0
Created	02/13/2025 13:55:11	**Observation Length**	56
Last Modified	02/13/2025 13:55:11	**Deleted Observations**	0
Protection		**Compressed**	NO
Data Set Type		**Sorted**	NO
Label			
Data Representation	SOLARIS_X86_64, LINUX_X86_64, ALPHA_TRU64, LINUX_IA64		
Encoding	utf-8 Unicode (UTF-8)		

Engine/Host Dependent Information	
Data Set Page Size	131072
Number of Data Set Pages	5
First Data Page	1
Max Obs per Page	2334
Obs in First Data Page	2289
Number of Data Set Repairs	0
Filename	/home/u64135306/sasuser.v94/Cours Variables Qualitatives/Exercice 1/panel95_table.sas7bdat
Release Created	9.0401M7
Host Created	Linux
Inode Number	2230767809
Access Permission	rw-r--r--
Owner Name	u64135306
File Size	768KB
File Size (bytes)	786432

Alphabetic List of Variables and Attributes

#	Variable	Type	Len	Format	Informat
5	etudes	Char	19	$19.	$19.
4	exper	Num	8	BEST12.	BEST32.
6	lw	Num	8	BEST12.	BEST32.
1	mident	Num	8	BEST12.	BEST32.
2	mois	Num	8	BEST12.	BEST32.
3	sexe	Char	5	$5.	$5.

Distribution du Log-Salaire (lw)

The UNIVARIATE Procedure
Variable: lw

Moments			
N	8856	**Sum Weights**	8856
Mean	3.88876114	**Sum Observations**	34438.8687
Std Deviation	0.46744023	**Variance**	0.21850037
Skewness	0.01185758	**Kurtosis**	1.27009394
Uncorrected SS	135859.355	**Corrected SS**	1934.82075
Coeff Variation	12.0202864	**Std Error Mean**	0.00496715

Basic Statistical Measures			
Location		**Variability**	
Mean	3.888761	**Std Deviation**	0.46744
Median	3.872802	**Variance**	0.21850
Mode	3.569616	**Range**	3.77177
		Interquartile Range	0.60614

Note: The mode displayed is the smallest of 3 modes with a count of 108.

Tests for Location: Mu0=0				
Test	**Statistic**		**p Value**	
Student's t	**t**	782.8958	**Pr > |t|**	<.0001
Sign	**M**	4428	**Pr >= |M|**	<.0001
Signed Rank	**S**	19609398	**Pr >= |S|**	<.0001

Quantiles (Definition 5)	
Level	**Quantile**
100% Max	5.53673
99%	5.14852
95%	4.63174
90%	4.46059
75% Q3	4.17575
50% Median	3.87280
25% Q1	3.56962
10%	3.36198
5%	3.21888
1%	2.64636

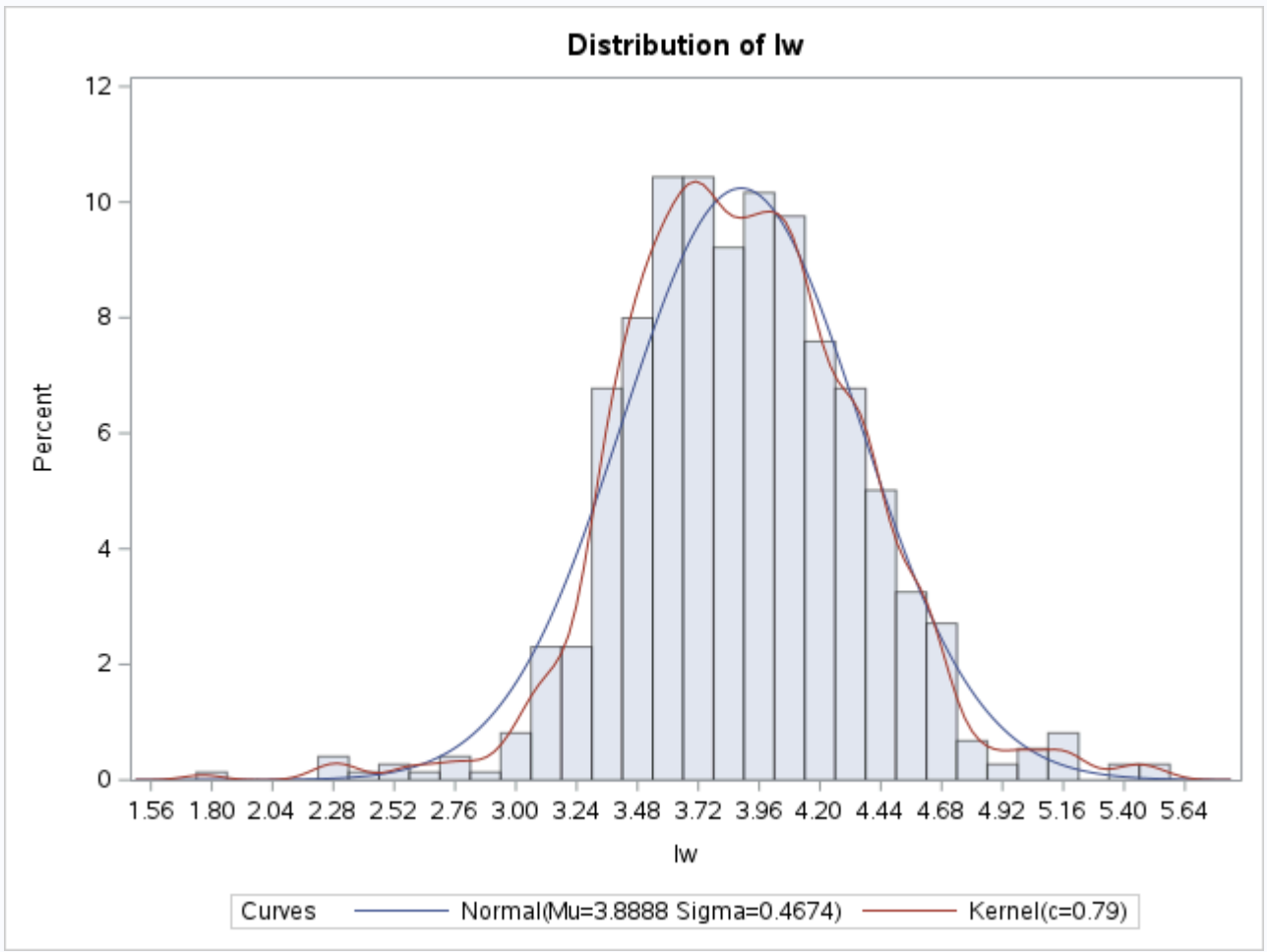
Quantiles (Definition 5)	
Level	Quantile
0% Min	1.76496

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
1.76496	4956	5.53673	3008
1.76496	4955	5.53673	3009
1.76496	4954	5.53673	3010
1.76496	4953	5.53673	3011
1.76496	4952	5.53673	3012

Missing Values			
Missing Value	Count	Percent Of	
		All Obs	Missing Obs
.	1692	16.04	100.00

Distribution du Log-Salaire (lw)

The UNIVARIATE Procedure



Distribution du Log-Salaire (lw)

The UNIVARIATE Procedure
Fitted Normal Distribution for lw

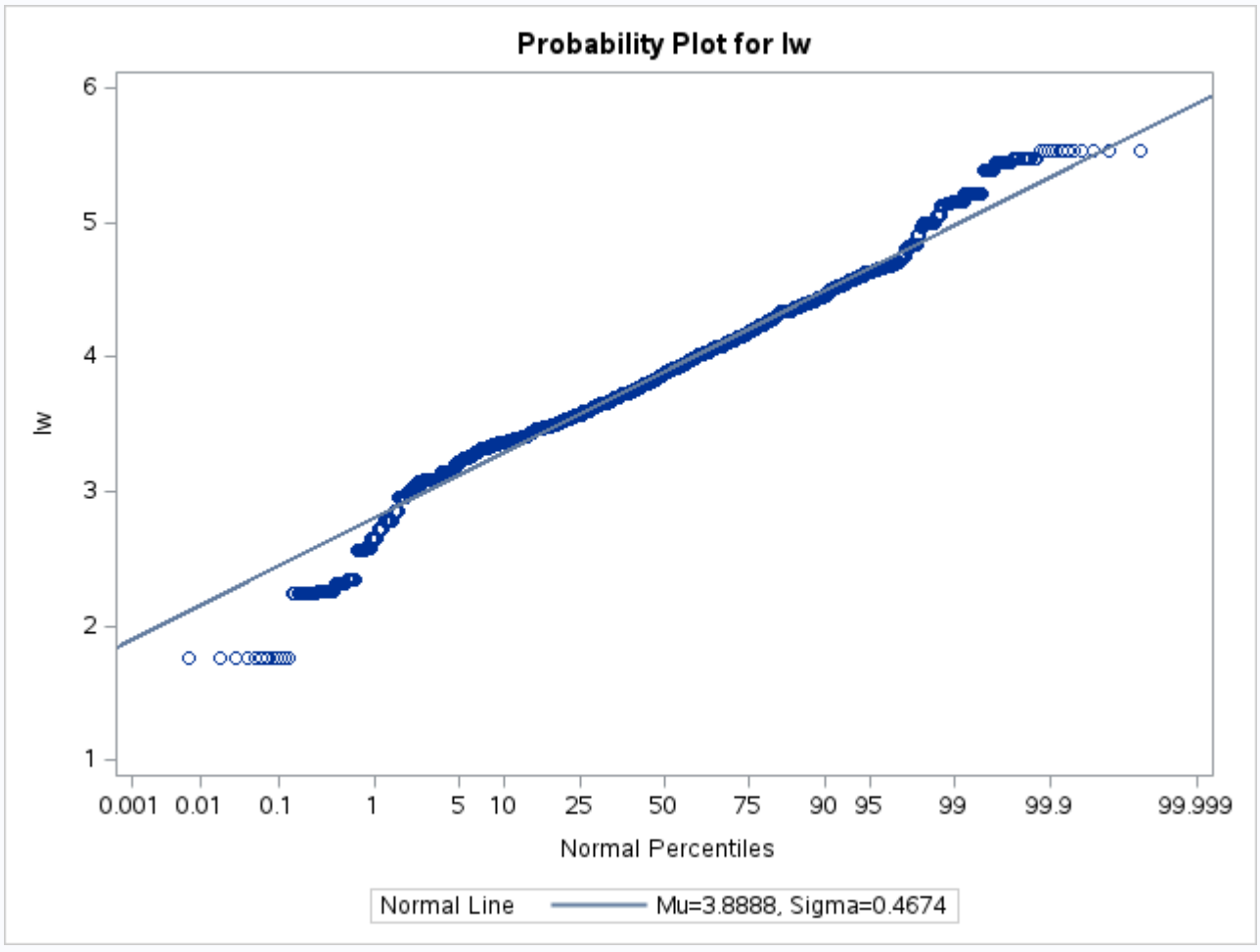
Parameters for Normal Distribution		
Parameter	Symbol	Estimate
Mean	Mu	3.888761
Std Dev	Sigma	0.46744

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.0405034	Pr > D	<0.010
Cramer-von Mises	W-Sq	1.8746503	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	17.2892902	Pr > A-Sq	<0.005

Quantiles for Normal Distribution		
Percent	Quantile	
	Observed	Estimated
1.0	2.64636	2.80133
5.0	3.21888	3.11989
10.0	3.36198	3.28971
25.0	3.56962	3.57348
50.0	3.87280	3.88876
75.0	4.17575	4.20404
90.0	4.46059	4.48781
95.0	4.63174	4.65763
99.0	5.14852	4.97619

Distribution du Log-Salaire (lw)

The UNIVARIATE Procedure



Distribution de l'expérience en Année

The UNIVARIATE Procedure
Variable: exper_years

Moments			
N	8124	Sum Weights	8124
Mean	16.1236973	Sum Observations	130988.917
Std Deviation	10.3007781	Variance	106.10603
Skewness	0.37692282	Kurtosis	-0.7639357
Uncorrected SS	2973924.92	Corrected SS	861899.285
Coeff Variation	63.885956	Std Error Mean	0.11428391

Basic Statistical Measures			
Location		Variability	
Mean	16.12370	Std Deviation	10.30078
Median	15.00000	Variance	106.10603
Mode	0.00000	Range	47.08333
		Interquartile Range	16.66667

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	141.0846	Pr > |t|	<.0001
Sign	M	3990	Pr >= |M|	<.0001
Signed Rank	S	15922095	Pr >= |S|	<.0001

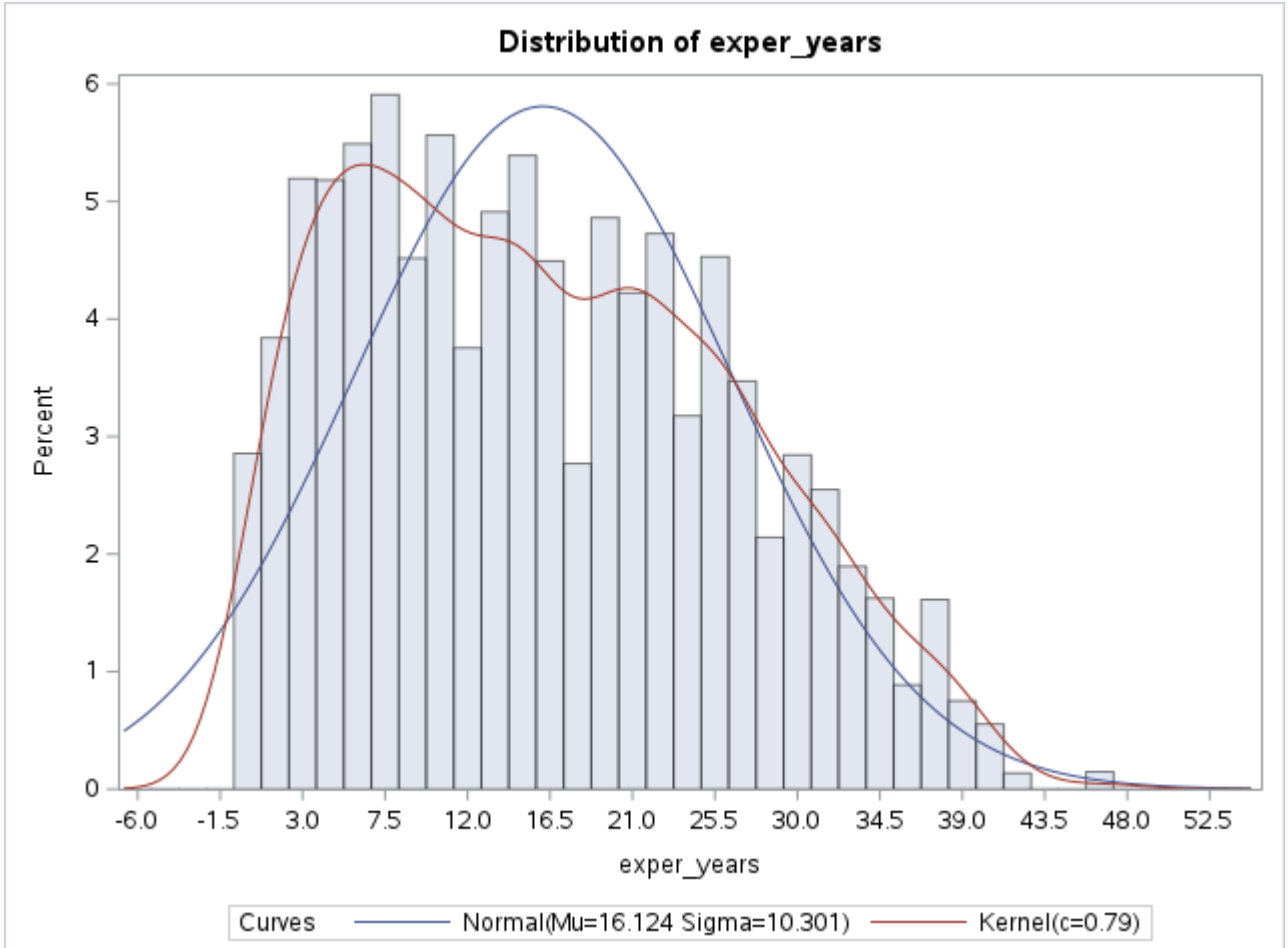
Quantiles (Definition 5)	
Level	Quantile
100% Max	47.08333
99%	39.41667
95%	34.41667
90%	30.75000
75% Q3	23.91667
50% Median	15.00000
25% Q1	7.25000
10%	3.08333
5%	1.75000
1%	0.00000
0% Min	0.00000

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
0	9984	46.7500	1520
0	9983	46.8333	1521
0	9982	46.9167	1522
0	9981	47.0000	1523
0	9980	47.0833	1524

Missing Values			
Missing Value	Count	Percent Of	
		All Obs	Missing Obs
.	2424	22.98	100.00

Distribution de l'expérience en Année

The UNIVARIATE Procedure



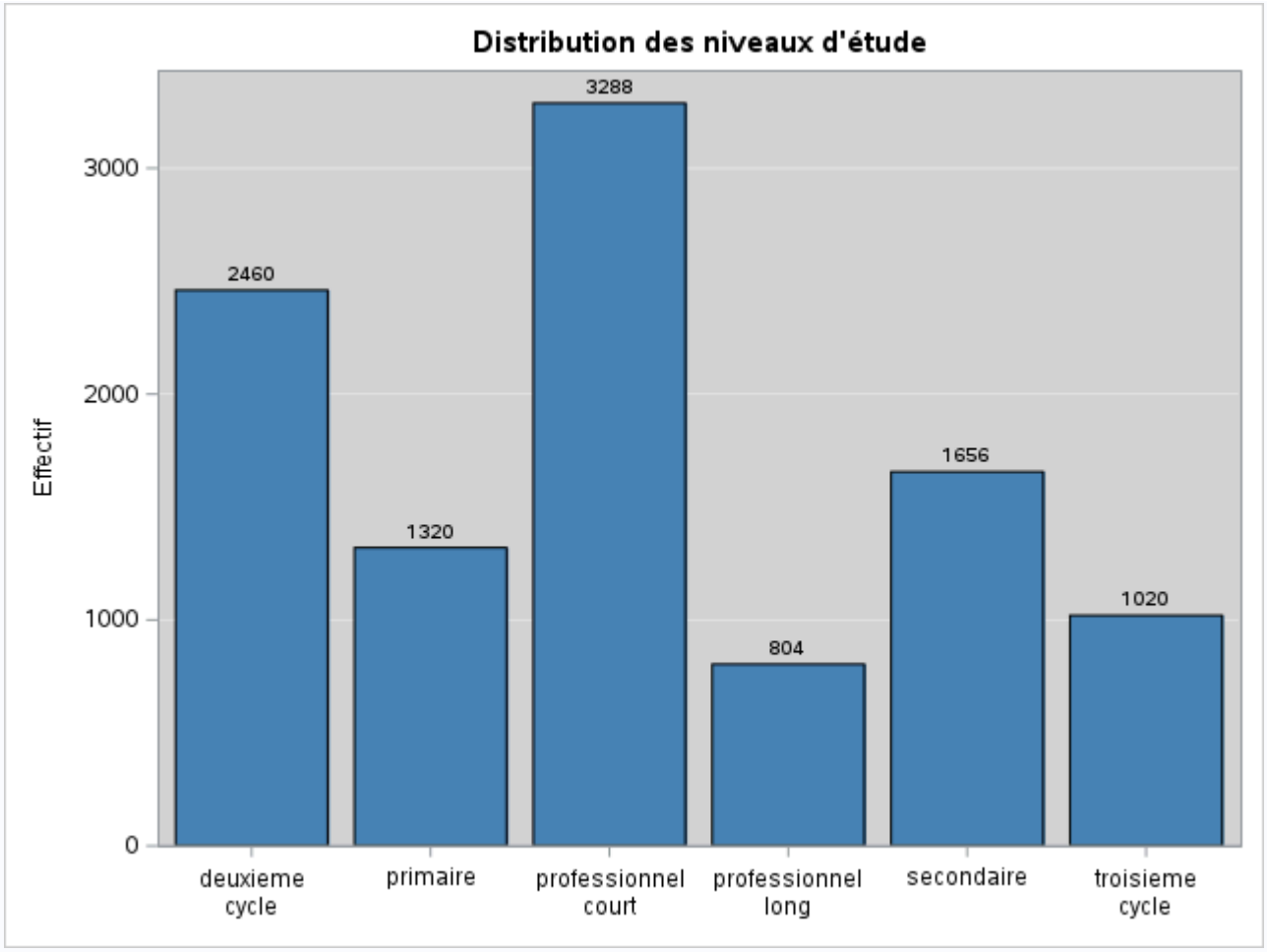
Distribution de l'expérience en Année

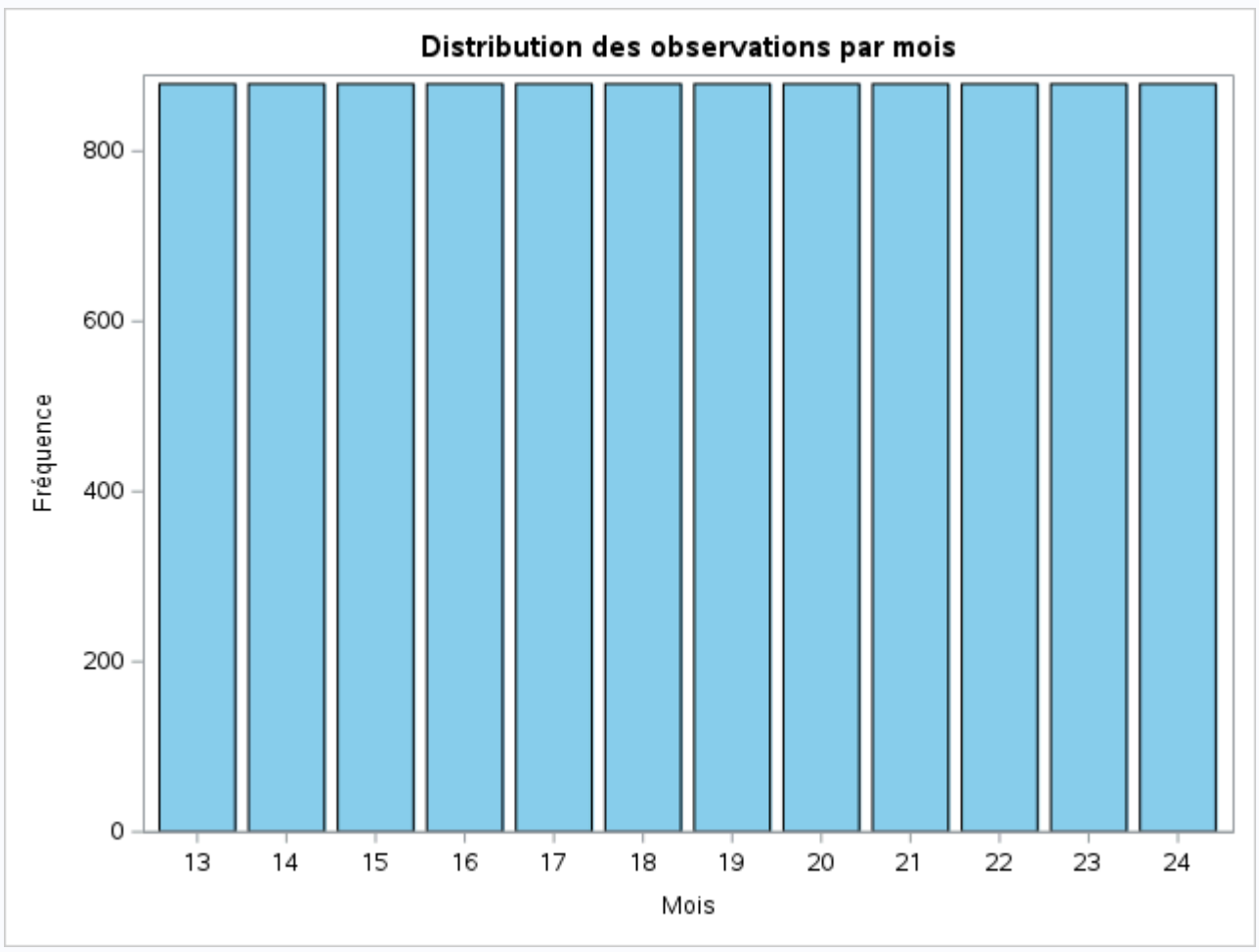
The UNIVARIATE Procedure
Fitted Normal Distribution for exper_years

Parameters for Normal Distribution		
Parameter	Symbol	Estimate
Mean	Mu	16.1237
Std Dev	Sigma	10.30078

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.0746718	Pr > D	<0.010
Cramer-von Mises	W-Sq	11.0878042	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	74.3875945	Pr > A-Sq	<0.005

Quantiles for Normal Distribution		
Percent	Quantile	
	Observed	Estimated
1.0	0.0000	-7.83950
5.0	1.7500	-0.81958
10.0	3.0833	2.92272
25.0	7.2500	9.17593
50.0	15.0000	16.12370
75.0	23.9167	23.07147
90.0	30.7500	29.32468
95.0	34.4167	33.06697
99.0	39.4167	40.08689

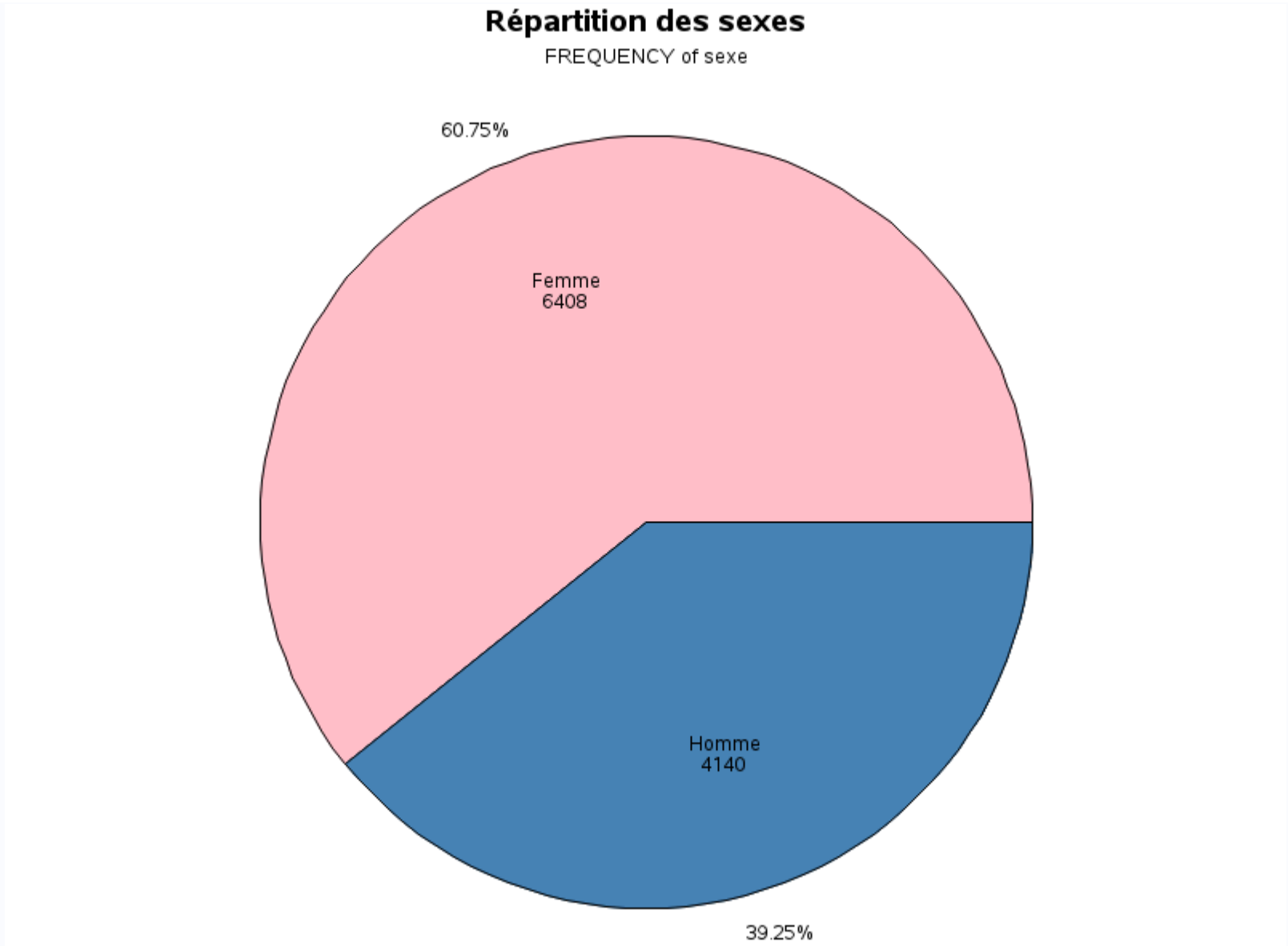




Distribution des observations par mois

The FREQ Procedure

sexe	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Femme	6408	60.75	6408	60.75
Homme	4140	39.25	10548	100.00



Répartition des sexes

The MEANS Procedure

Analysis Variable : lw						
etudes	**N Obs**	**N**	**Mean**	**Std Dev**	**Minimum**	**Maximum**
deuxieme cycle	2460	2292	4.0704467	0.4395550	1.7649612	5.3873510
primaire	1320	924	3.6030209	0.2963112	2.6463566	4.2738128
professionnel court	3288	2664	3.7261213	0.3681880	2.3494897	4.9871631
professionnel long	804	732	3.8576809	0.3645359	2.7835577	4.9991741
secondaire	1656	1308	3.8352762	0.4889569	2.2572458	5.2175918
troisieme cycle	1020	936	4.2878874	0.5327744	2.2478602	5.5367284

Répartition des sexes

The MEANS Procedure

Analysis Variable : lw				
N	**Mean**	**Std Dev**	**Minimum**	**Maximum**
8856	3.8887611	0.4674402	1.7649612	5.5367284

Régression du log-salaire sur tous les niveaux d'études (sans précaution)

The REG Procedure
Model: MODEL1
Dependent Variable: lw

Number of Observations Read	10548
Number of Observations Used	8856
Number of Observations with Missing Values	1692

Analysis of Variance					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
Model	5	375.12300	75.02460	425.70	<.0001
Error	8850	1559.69775	0.17624		
Corrected Total	8855	1934.82075			

Root MSE	0.41981	**R-Square**	0.1939
Dependent Mean	3.88876	**Adj R-Sq**	0.1934
Coeff Var	10.79536		

Note: Model is not full rank. Least-squares solutions for the parameters are not unique. Some statistics will be misleading. A reported DF of 0 or B means that the estimate is biased.

Note: The following parameters have been set to 0, since the variables are a linear combination of other variables as shown.

Cycle3 =	Intercept - Primaire - Secondaire - Professionnel_court - Professionnel_long - Cycle2

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	B	4.28789	0.01372	312.49	<.0001
Primaire	B	-0.68487	0.01947	-35.18	<.0001
Secondaire	B	-0.45261	0.01797	-25.18	<.0001
Professionnel_court	B	-0.56177	0.01595	-35.22	<.0001
Professionnel_long	B	-0.43021	0.02071	-20.77	<.0001
Cycle2	B	-0.21744	0.01628	-13.35	<.0001
Cycle3	0	0	.	.	.

Régression du log-salaire sur tous les niveaux d'études (sans constante)

The REG Procedure
Model: MODEL1
Dependent Variable: lw

Number of Observations Read	10548
Number of Observations Used	8856
Number of Observations with Missing Values	1692

Note: No intercept in model. R-Square is redefined.

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	134300	22383	127007	<.0001
Error	8850	1559.69775	0.17624		
Uncorrected Total	8856	135859			

Root MSE	0.41981	R-Square	0.9885
Dependent Mean	3.88876	Adj R-Sq	0.9885
Coeff Var	10.79536		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Primaire	1	3.60302	0.01381	260.89	<.0001
Secondaire	1	3.83528	0.01161	330.41	<.0001
Professionnel_court	1	3.72612	0.00813	458.12	<.0001
Professionnel_long	1	3.85768	0.01552	248.62	<.0001
Cycle2	1	4.07045	0.00877	464.20	<.0001
Cycle3	1	4.28789	0.01372	312.49	<.0001

Régression du log-salaire en prenant 'Primaire' comme référence

The REG Procedure
Model: MODEL1
Dependent Variable: lw

Number of Observations Read	10548
Number of Observations Used	8856
Number of Observations with Missing Values	1692

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	375.12300	75.02460	425.70	<.0001
Error	8850	1559.69775	0.17624		
Corrected Total	8855	1934.82075			

Root MSE	0.41981	R-Square	0.1939
Dependent Mean	3.88876	Adj R-Sq	0.1934
Coeff Var	10.79536		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	3.60302	0.01381	260.89	<.0001
Secondaire	1	0.23226	0.01804	12.87	<.0001
Professionnel_court	1	0.12310	0.01603	7.68	<.0001
Professionnel_long	1	0.25466	0.02077	12.26	<.0001
Cycle2	1	0.46743	0.01636	28.57	<.0001
Cycle3	1	0.68487	0.01947	35.18	<.0001

Régression du log-salaire en prenant 'cycle3' comme référence

The REG Procedure
Model: MODEL1
Dependent Variable: lw

Number of Observations Read	10548
Number of Observations Used	8856
Number of Observations with Missing Values	1692

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	375.12300	75.02460	425.70	<.0001
Error	8850	1559.69775	0.17624		
Corrected Total	8855	1934.82075			

Root MSE	0.41981	R-Square	0.1939
Dependent Mean	3.88876	Adj R-Sq	0.1934
Coeff Var	10.79536		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	4.28789	0.01372	312.49	<.0001
Primaire	1	-0.68487	0.01947	-35.18	<.0001
Secondaire	1	-0.45261	0.01797	-25.18	<.0001
Professionnel_court	1	-0.56177	0.01595	-35.22	<.0001
Professionnel_long	1	-0.43021	0.02071	-20.77	<.0001
Cycle2	1	-0.21744	0.01628	-13.35	<.0001

Régression avec contrainte sur la moyenne des coefficients

The REG Procedure
Model: MODEL1
Dependent Variable: lw

Note: Restrictions have been applied to parameter estimates.

Number of Observations Read	10548
Number of Observations Used	8856
Number of Observations with Missing Values	1692

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	375.12300	75.02460	425.70	<.0001
Error	8850	1559.69775	0.17624		
Corrected Total	8855	1934.82075			

Root MSE	0.41981	R-Square	0.1939
Dependent Mean	3.88876	Adj R-Sq	0.1934
Coeff Var	10.79536		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	3.88876	0.00446	871.73	<.0001
Primaire	1	-0.28574	0.01307	-21.86	<.0001
Secondaire	1	-0.05348	0.01072	-4.99	<.0001
Professionnel_court	1	-0.16264	0.00680	-23.91	<.0001
Professionnel_long	1	-0.03108	0.01486	-2.09	0.0365
Cycle2	1	0.18169	0.00755	24.07	<.0001
Cycle3	1	0.39913	0.01298	30.76	<.0001
RESTRICT	-1	6.46492E-14	3.94327E-11	0.00	0.9987*

*** Probability computed using beta distribution.**

Régression avec contrainte sur la moyenne NON pondérée des coefficients

The REG Procedure
Model: MODEL1
Dependent Variable: lw

Note: Restrictions have been applied to parameter estimates.

Number of Observations Read	10548
Number of Observations Used	8856
Number of Observations with Missing Values	1692

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	375.12300	75.02460	425.70	<.0001
Error	8850	1559.69775	0.17624		
Corrected Total	8855	1934.82075			

Root MSE	0.41981	R-Square	0.1939
Dependent Mean	3.88876	Adj R-Sq	0.1934
Coeff Var	10.79536		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	3.89674	0.00499	780.40	<.0001
Primaire	1	-0.29372	0.01233	-23.82	<.0001
Secondaire	1	-0.06146	0.01071	-5.74	<.0001
Professionnel_court	1	-0.17062	0.00831	-20.53	<.0001
Professionnel_long	1	-0.03906	0.01362	-2.87	0.0041
Cycle2	1	0.17371	0.00873	19.90	<.0001
Cycle3	1	0.39115	0.01227	31.89	<.0001
RESTRICT	-1	9.54264E-11	5.692283E-8	0.00	0.9987*

*** Probability computed using beta distribution.**

Corrélation entre le log-salaire observé et les log-salaires prédits

The CORR Procedure

7 Variables:	lw lw_pred_A lw_pred_B lw_pred_C lw_pred_D lw_pred_E lw_pred_F

Simple Statistics							
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum	Label
lw	8856	3.88876	0.46744	34439	1.76496	5.53673	
lw_pred_A	10548	3.87251	0.20452	40847	3.60302	4.28789	Predicted Value of lw
lw_pred_B	10548	3.87251	0.20452	40847	3.60302	4.28789	Predicted Value of lw
lw_pred_C	10548	3.87251	0.20452	40847	3.60302	4.28789	Predicted Value of lw
lw_pred_D	10548	3.87251	0.20452	40847	3.60302	4.28789	Predicted Value of lw
lw_pred_E	10548	3.87251	0.20452	40847	3.60302	4.28789	Predicted Value of lw
lw_pred_F	10548	3.87251	0.20452	40847	3.60302	4.28789	Predicted Value of lw

Pearson Correlation Coefficients Prob > |r| under H0: Rho=0 Number of Observations							
	lw	lw_pred_A	lw_pred_B	lw_pred_C	lw_pred_D	lw_pred_E	lw_pred_F
lw	1.00000 8856	0.44032 <.0001 8856	0.44032 <.0001 8856	0.44032 <.0001 8856	0.44032 <.0001 8856	0.44032 <.0001 8856	0.44032 <.0001 8856
lw_pred_A Predicted Value of lw	0.44032 <.0001 8856	1.00000 10548	1.00000 <.0001 10548	1.00000 <.0001 10548	1.00000 <.0001 10548	1.00000 <.0001 10548	1.00000 <.0001 10548
lw_pred_B Predicted Value of lw	0.44032 <.0001 8856	1.00000 <.0001 10548	1.00000 10548	1.00000 <.0001 10548	1.00000 <.0001 10548	1.00000 <.0001 10548	1.00000 <.0001 10548
lw_pred_C Predicted Value of lw	0.44032 <.0001 8856	1.00000 <.0001 10548	1.00000 <.0001 10548	1.00000 10548	1.00000 <.0001 10548	1.00000 <.0001 10548	1.00000 <.0001 10548
lw_pred_D Predicted Value of lw	0.44032 <.0001 8856	1.00000 <.0001 10548	1.00000 <.0001 10548	1.00000 <.0001 10548	1.00000 10548	1.00000 <.0001 10548	1.00000 <.0001 10548
lw_pred_E Predicted Value of lw	0.44032 <.0001 8856	1.00000 <.0001 10548	1.00000 <.0001 10548	1.00000 <.0001 10548	1.00000 <.0001 10548	1.00000 10548	1.00000 <.0001 10548
lw_pred_F Predicted Value of lw	0.44032 <.0001 8856	1.00000 <.0001 10548	1.00000 <.0001 10548	1.00000 <.0001 10548	1.00000 <.0001 10548	1.00000 <.0001 10548	1.00000 10548

Synthèse des résultats des régressions

Model	Root_MSE	Dep_Mean	R2	Adj_R2	Intercept	Primaire	Secondaire	Professionnel_court	Professionnel_long	Cycle2	Cycle3
A: Tous les niveaux d'étude, 'sans précaution'	0.419806	3.888761	0.419806	3.888761	4.28789	-0.68487	-0.45261	-0.56177	-0.43021	-0.21744	0
B: Tous les niveaux d'étude, sans constante	0.419806	3.888761	0.419806	3.888761	.	3.60302	3.83528	3.72612	3.85768	4.07045	4.28789
C: Tous les niveaux d'étude sauf 'primaire'	0.419806	3.888761	0.419806	3.888761	3.60302	.	0.23226	0.12310	0.25466	0.46743	0.68487
D: Tous les niveaux d'étude sauf 'cycle3'	0.419806	3.888761	0.419806	3.888761	4.28789	-0.68487	-0.45261	-0.56177	-0.43021	-0.21744	.
E: Nullité de la moyenne des coefs pondérée par les effectifs	0.419806	3.888761	0.419806	3.888761	3.88876	-0.28574	-0.05348	-0.16264	-0.03108	0.18169	0.39913
F: Nullité de la la moyenne NON pondérée	0.419806	3.888761	0.419806	3.888761	3.89674	-0.29372	-0.06146	-0.17062	-0.03906	0.17371	0.39115