STA380.18 – Homework 2

Principal Components Analysis

– **Avani Sharma (as85253)**

1. Create a SAS dataset called **WORK.RATINGS** that contains the data in the **job ratings.txt** file. Assign the SAS names **JOB, KNOWHOW, PROBLEM_SOLVING, ACCOUNTABILITY, SALARY**, respectively, to the five variables as they appear from left to right in the file. Extract the principal components of the three dimensions that were rated by the management consulting firm. Use the default (standardized) version of the extraction. *Your answer for question 1 is your SAS code only.*

1:

```
**Q1.;
data WORK.RATINGS;
        input job knowhow problem_solving accountability salary;
        cards;

0  800  608  1056  102000
2  528  304  460  75740
3  460  264  460  75740
5  528  304  304  79172
4  460  264  400  70000
0  460  264  400  66536
0  528  304  264  70000
7  460  230  264  68000
10  400  200  350  73140
7  400  175  230  66016
7  400  200  200  66016
5  400  175  200  71840
...
...
Run;

**Principal Components;

PROC Princomp data= WORK.RATINGS;
var knowhow problem_solving accountability;
RUN;
```

2. This question verifies the basic property of principal components transformations.

   a. Write the equations of the principal components of the PCA in question 1.
   b. Verify that the principal component transformation in question 1 is an orthonormal rotation of the (standardized) original three dimensions by showing that the rotation matrix satisfies the definition of an orthonormal transformation.

2: a. Equations of the principal components of PCA in question 1 are -

PC 1 = 0.576251 * knowhow + 0.584343 * problem_solving + 0.571383 * accountability
PC 2 = -0.618121 * knowhow + -0.145758 * problem_solving + 0.772451 * accountability
PC 3 = 0.534660 * knowhow + -0.798310 * problem_solving + 0.277201 * accountability

b. To prove that principal component transformation is an orthogonal rotation of the original 3 dimensions, two conditions should be true : Vectors should be of unit length & orthogonal to each other –

i. 3 Vectors (principal components) must be of unit length

Length (PC 1) = 0.576251 * 0.576251 + 0.584343 * 0.584343 + 0.571383 * 0.571383 = 1
Length (PC 2) = -0.618121 * -0.618121 + -0.145758 * -0.145758 + 0.772451 * 0.772451 = 1
Length (PC 3) = 0.534660 * 0.534660 + -0.798310 * -0.798310 + 0.277201 * 0.277201 = 1

ii. 3 Vectors (principal components) must be orthogonal to each other (i.e. dot product of any 2 pair of vectors must be zero)

Dot Products:
PC 1 and PC 2 = 0.576251 * -0.618121 + 0.584343 * -0.145758 + 0.571383 * 0.772451 = 0
PC 2 and PC 3 = -0.618121 * 0.534660 + -0.145758 * -0.798310 + 0.772451 * 0.277201 = 0
PC 3 and PC 1 = 0.534660 * 0.576251 + -0.798310 * 0.584343 + 0.277201 * 0.571383 = 0

On Excel:

Unit Vectors:

	Eigenvectors			Eigenvector Squares			
	Prin1	Prin2	Prin3	Prin1*Prin1	Prin2*Prin2	Prin3*Prin3	Sum
knowhow	0.576251	-0.61812	0.53466	0.332065215	0.382073571	0.285861316	1
problem_solving	0.584343	-0.14576	-0.79831	0.341456742	0.021245395	0.637298856	1
accountability	0.571383	0.772451	0.277201	0.326478533	0.596680547	0.076840394	1

Dot Product:

	Dot Product			Sum
Prin1.Prin2	0.336728	0.090096	-0.42682	0
Prin2.Prin3	0.333884	-0.11259	-0.22129	0
Prin3.Prin1	0.32926	-0.47747	0.148208	0

From the above calculations, the first & second conditions are met. Thus, principal component transformation is an orthonormal transformation.

3. This question partially verifies the geometry-preserving property of principal components transformations.

a. Rotate the first two jobs in the text file by calculating their principal component scores.
b. The rotated scores for the two jobs in part (a) are each a vector of three scores. Verify that the lengths of these two vectors are the same as the lengths of the original (but standardized) ratings vectors of the two jobs.
c. Verify that the angle between these two rotated vectors is the same as the angle between the original unrotated vectors.

3:

a. Principal Components for the first 2 jobs in the text file are shown below –

```
** Q3;
PROC princomp data = WORK.RATINGS OUT= WORK.RATINGS_PCA;
var knowhow problem_solving accountability;
RUN;
```

From SAS:

job	Prin1	Prin2	Prin3
0	9.089332156	1.2654300739	-0.19679536
2	3.7513631804	-0.059567149	0.0985589143
3	3.2058572773	0.3254455567	0.1501086816
5	3.1543849719	-0.866619424	-0.191059403

On Excel:

For Instance:

	Prin1
knowhow	0.576251
problem_solving	0.584343
accountability	0.571383

Prin 1:
standardised_knowhow*0.576+standardised_problemsolving*0.584+standardised_accountability*0.571

Similar calculations Applied,

Standardised Xs				Calculated Principal Components		
job	knowhow	problem_solving	accountability	Prin1	Prin2	Prin3
0	4.35032492	5.283939792	6.116424944	9.089334	1.265435	-0.1968
2	2.25124045	2.122082877	2.124774933	3.751364	-0.05957	0.098558

b. Verifying that the lengths of the standardized vector and PCs is same

Standardised Xs				Calculated Principal Components			Original Length	Transformed Length
job	knowhow	problem_solving	accountability	Prin1	Prin2	Prin3		
0	4.350325	5.283939792	6.116424944	9.089334	1.265435	0.196798143	9.1791068	9.17910885
2	2.25124	2.122082877	2.124774933	3.751364	-0.05957	0.098557974	3.7531304	3.75313127

c. In order to verify that the angle between the 2 rotated vectors and the original vectors is same, we will calculate cosine of the angle between the 2 vectors.

Cosine (angle) between 2 vectors = dot product (vec 1, vec 2) / length (vec 1) * length (vec 2)

Cosine (angle) between original vectors	0.987002391
Cosine (angle) between rotated vectors	0.9870024

The cosine of the angle between 2 original vectors is same as in the rotated vectors

4. Obtain the principal components scores for all 67 jobs. Calculate the variances of the three sets of scores and verify that the variances are equal to the eigenvalues of the PC transformation.

4)

The eigen values of the PC transformation are shown below –

Eigenvalues of the Correlation Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	2.90808114	2.82438377	0.9694	0.9694
2	0.08369737	0.07547588	0.0279	0.9973
3	0.00822149		0.0027	1.0000

The variance of the principal component score across all the 67 jobs are **2.90808**, **0.08370** and **0.00822** as obtained from the calculations done in excel. Clearly, variances are equal to the eigenvalues of PCs.

Standardised kh	Standardised ps	Standardised acc	Prin1	Prin2	Prin3
4.35	5.28	6.12	9.09	1.27	-0.20
2.25	2.12	2.12	3.75	-0.06	0.10
1.73	1.71	2.12	3.21	0.33	0.15
2.25	2.12	1.08	3.15	-0.87	-0.19
1.73	1.71	1.72	2.98	0.02	0.04
1.73	1.71	1.72	2.98	0.02	0.04
2.25	2.12	0.81	3.00	-1.07	-0.27
1.73	1.35	0.81	2.25	-0.64	0.07
1.26	1.04	1.39	2.13	0.14	0.23
1.26	0.78	0.58	1.52	-0.44	0.21
1.26	1.04	0.38	1.56	-0.64	-0.05
1.26	0.78	0.38	1.40	-0.60	0.16
0.52	0.16	0.22	0.52	-0.18	0.21
0.21	0.00	0.22	0.25	0.03	0.17
0.21	0.00	0.22	0.25	0.03	0.17
-0.05	0.00	-0.07	-0.07	-0.03	-0.05

	Variances
Prin1	2.90808
Prin2	0.08370
Prin3	0.00822

5. Find the regression equation that results from regressing **PRIN1** on the three ratings knowhow, problem_solving, and accountability after the ratings have been standardized and without an intercept.2 Are you surprised by the equation?

5)

```
** Q5;

PROC STDIZE DATA=WORK.RATINGS OUT=WORK.STD;
     VAR knowhow problem_solving accountability;

PROC princomp data = WORK.STD OUT= WORK.STD_PCA;
     var knowhow problem_solving accountability;
     RUN;

PROC REG DATA = WORK.STD_PCA;
     model Prin1 = knowhow problem_solving accountability / noint;
     RUN;
```

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
knowhow	1	0.57625	0	Infty	<.0001
problem_solving	1	0.58434	0	Infty	<.0001
accountability	1	0.57138	0	Infty	<.0001

The regression equation is as follows.

Prin1 = 0.57625 * knowhow + 0.58434 * problem_solving + 0.57138 * accountability

The regression coefficients are equal to the loadings that we obtained for Prin1 in question 2, part b. The prinicipal components are orthogonal unit vectors which contain the entire information conveyed by the variables. So, for each variable when we regress it with principal components, regressing it alone or in group would give the same coeffcients (no multicollineraity). This is not that surprising, since Principal components are a linear combination of the variables.

6. Find the regression equation that results from regressing (standardized) **KNOWHOW** on the three principal components without an intercept. Are you surprised by the equation?

6)

```
** Q6;

PROC REG DATA = WORK.STD_PCA;
    model knowhow = Prin1 Prin2 Prin3 / noint;
    RUN;
```

The results from SAS:

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Prin1	1	0.57625	0	Infty	<.0001
Prin2	1	-0.61812	0	-Infty	<.0001
Prin3	1	0.53466	0	Infty	<.0001

From questin 2 calculations:

	Eigenvectors			Eigenvector Squares			
	Prin1	Prin2	Prin3	Prin1*Prin1	Prin2*Prin2	Prin3*Prin3	Sum
knowhow	0.576251	-0.61812	0.53466	0.332065215	0.382073571	0.285861316	1
problem_solving	0.584343	-0.14576	-0.79831	0.341456742	0.021245395	0.637298856	1
accountability	0.571383	0.772451	0.277201	0.326478533	0.596680547	0.076840394	1

The regression equation is as follows.

Knowhow = 0.57625 * Prin1 – 0.61812 * Prin2 + 0.53466 * Prin3

The coefficients are just the loadings that we obtained for `knowhow` when we ran a principal component analysis in question 1 (image picked from 2). The results are not surprising because original variables are also a linear combination of principal components. These PCs are orthogonal to each other and hence aren't multicollinear, running regression individually or with all PCs at a time would yield the same coefficients.

7. Write the **loadings matrix**, structured with components as columns and variables as rows. Using the loadings matrix, try to interpret meanings for the three principal components.

7) The loading matrix from SAS program:

Eigenvectors			
	Prin1	Prin2	Prin3
knowhow	0.576251	-.618121	0.534660
problem_solving	0.584343	-.145758	-.798310
accountability	0.571383	0.772451	0.277201

not loadings matrix

-5

Interpretation –

1. **Prin1**: Loadings for all three variables are equally high for the first principal component implying an all rounded score for all three traits. Such can be the requirement of jobs with holiistic responsibilities starting from inception of an idea to implementation. It could represent job profiles which are **manegerial or project owner level.**
2. **Prin2**: This component is positively high on accountability but negatively loaded with knowhow (problem-solving to a small extent). These can be **entry-level positions** which are focussed highly on execution or **assiting job profiles** like personal assisstants.
3. **Prin3**: This component has a high negative loading of problem-solving and smaller positive loading of knowhow, they can be associated with job profiles requiring **research, sales** which are required to know about the projects in and out but real world implementation is taken care of by others.

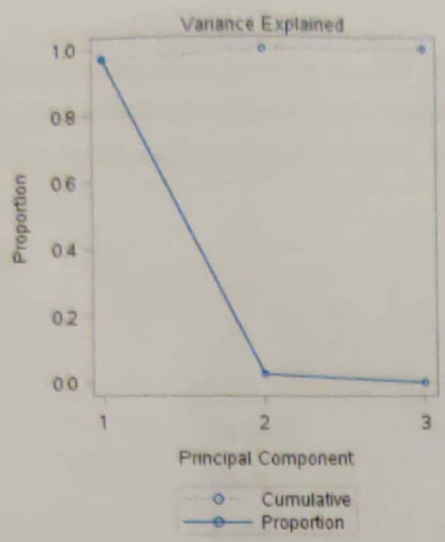
8. How many principal components would you retain ...

8) a. Kaiser Rule: The Kaiser rule is to drop all components with eigenvalues under 1.0 – this being the eigenvalue equal to the information accounted for by an average single item.. Thus we will only retain **one PC.**

b. Joliffe Rule: Retain all components that have eigen vale > 0.7. We will **retain 1 PC**

c. Using 80% rule – Again we will retain the **first PC.** (from scree plot)

Eigenvalues of the Correlation Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	2.90808114	2.82438377	0.9694	0.9694
2	0.08369737	0.07547588	0.0279	0.9973
3	0.00822149		0.0027	1.0000



9. Find the regression equation that results from regressing **salary** on the three principal components with intercept. How much explanatory power do the three PCs collectively have in explaining **salary**?

9)

```
** Q9;

PROC REG DATA = WORK. STD_PCA;
    model salary = Prin1 Prin2 Prin3;
RUN;
```

Regression equation of salary vs the 3 PC's is as follows, they have a collective power of explaining 90.03% of the variance in salary.

Salary = 63929 + 3557.20641* Prin1 + 2316.12408 * Prin2 + 3540.61136 * Prin3

Root MSE	2082.09165	**R-Square**	0.9003
Dependent Mean	63929	**Adj R-Sq**	0.8955
Coeff Var	3.25686		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	63929	254.36798	251.33	<.0001
Prin1	1	3557.20641	150.28811	23.67	<.0001
Prin2	1	2316.12408	885.87403	2.61	0.0112
Prin3	1	3540.61136	2826.52316	1.25	0.2150

10. In terms of explaining **salary**...
a) Which component is most useful? Second most useful? Least useful?
b) Is the usefulness of the PCs for explaining salary in the order PC1 > PC2 > PC3?
c) How much explanatory power is lost if one uses only PRIN1 to explain **salary**?

10)

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	2465105931	821701977	189.55	<.0001
Error	63	273111655	4335106		
Corrected Total	66	2738217587			

Root MSE	2082.09165	R-Square	0.9003
Dependent Mean	63929	Adj R-Sq	0.8955
Coeff Var	3.25686		

Parameter Estimates								
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Type I SS	Type II SS	Variance Inflation
Intercept	1	63929	254.36798	251.33	<.0001	2.738263E11	2.738263E11	0
Prin1	1	3557.20641	150.28811	23.67	<.0001	2428670447	2428670447	1.00000
Prin2	1	2316.12408	885.87403	2.61	0.0112	29633257	29633257	1.00000
Prin3	1	3540.61136	2826.52316	1.25	0.2150	6802227	6802227	1.00000

a. Type I SS value for Prin1 is maximum, it is the most important component. Prin2 is the second most useful and Prin3 is the least useful component
b. Yes, the usefulness of PCs in explaining salary is of the order PC1 > PC2 > PC3
c. Using only Prin1, we are able to explain the following amount of variance in salary : 0.9* 2428670447 / (2428670447 + 29633257 + 6802227) = 88.7%. Thus we have lost 90 – 88.7 = 1.3% explanatory power if we only use Prin1. VIFs here are all 1 as PCs are not multicollinear.

In statistics, the **variance inflation factor (VIF)** is the ratio of variance in a model with multiple terms, divided by the variance of a model with one term alone. It quantifies the severity of multicollinearity in an ordinary least squares regression analysis. It provides an index that measures how much the variance (the square of the estimate's standard deviation) of an estimated regression coefficient is increased because of collinearity. VIF < 10. The square root of the variance inflation factor indicates how much larger the standard error is, compared with what it would be if that variable were uncorrelated with the other predictor variables in the model.

CODE

```
** Q1;

data WORK.RATINGS;
     input job knowhow problem_solving accountability salary;
     cards;
0  800  608  1056  102000
2  528  304  460  75740
3  460  264  460  75740
5  528  304  304  79172
4  460  264  400  70000
0  460  264  400  66536
0  528  304  264  70000
7  460  230  264  68000
10  400  200  350  73140
7  400  175  230  66016
7  400  200  200  66016
5  400  175  200  71840
5  304  115  175  71580
2  264  100  175  65860
3  264  100  175  66432
10  230  100  132  64040
10  230  100  132  62610
7  230  87  132  65002
7  230  76  115  64001
5  230  76  115  66900
5  230  87  100  63000
5  230  87  100  63780
7  200  87  100  62000
7  200  76  100  61960
7  200  76  100  62012
7  200  76  87  62300
5  200  76  87  61960
7  200  66  87  61700
7  175  66  100  61440
2  175  57  100  62220
3  175  57  100  63260
7  175  57  100  59880
2  175  57  100  62480
3  175  57  100  63000
2  175  57  100  63260
3  175  57  100  62480
4  175  57  87  62480
7  175  57  87  61440
2  175  57  87  62064
3  175  57  87  61180
2  175  57  87  59100
3  175  57  87  59620
```

```
5  175  66  76  59880
5  175  66  76  60200
7  175  57  76  60140
7  175  57  76  61700
5  175  66  66  60000
7  152  50  87  60920
7  152  50  76  59100
3  152  50  76  61700
2  152  50  76  59880
3  152  50  76  61700
5  152  50  66  59360
5  152  43  66  60660
2  152  43  66  59984
2  152  43  66  60660
3  152  43  66  60920
3  152  43  66  60920
2  152  43  66  60920
3  152  43  66  60660
3  152  43  66  60660
7  152  43  66  58320
5  152  43  66  59360
2  152  43  66  60920
3  152  43  66  60920
4  152  43  66  60660
7  152  43  57  59880
RUN;

**Principal Components;

PROC Princomp data= WORK.RATINGS;
     var knowhow problem_solving accountability;
     RUN;

** Q5;

PROC STDIZE DATA=WORK.RATINGS OUT=WORK.STD;
     VAR knowhow problem_solving accountability;

PROC princomp data = WORK.STD OUT= WORK.STD_PCA;

     var knowhow problem_solving accountability;
     RUN;


PROC REG DATA = WORK.STD_PCA;
     model Prin1 = knowhow problem_solving accountability / noint;
     RUN;
```

```
** Q6;
PROC REG DATA=Work.STD_PCA;
     model knowhow = Prin1 Prin2 Prin3 / noint;
     RUN;

** Q9;

PROC REG DATA = WORK.STD_PCA;
     model salary = Prin1 Prin2 Prin3;
     RUN;

** Q10;
proc reg DATA=WORK.STd_PCA;
  model salary = PRIN1-PRIN3 / ss1 ss2 vif;
  RUN;
```

The PRINCOMP Procedure

Observations	67
Variables	3

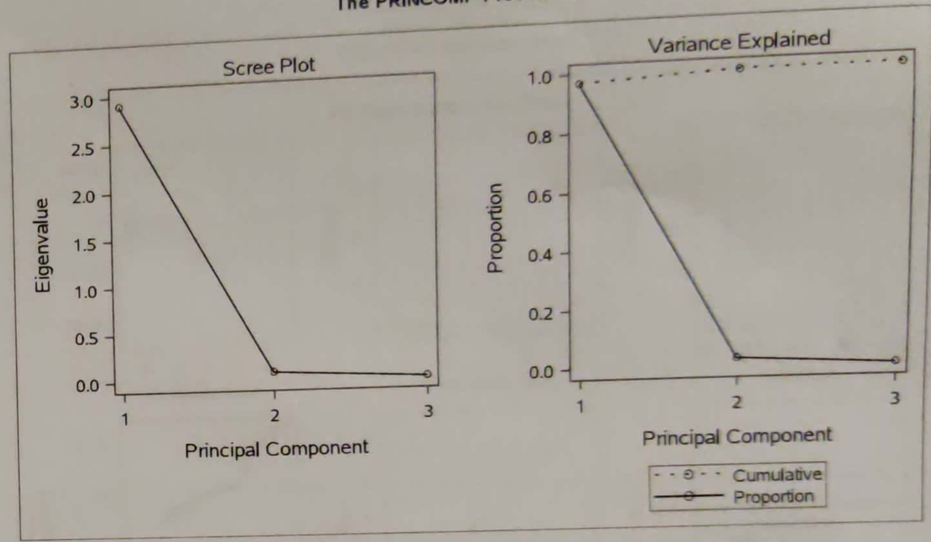
Simple Statistics			
	knowhow	problem_solving	accountability
Mean	236.2835821	99.97014925	142.7462687
StD	129.5803022	96.14603321	149.3116878

Correlation Matrix			
	knowhow	problem_solving	accountability
knowhow	1.0000	0.9833	0.9188
problem_solving	0.9833	1.0000	0.9597
accountability	0.9188	0.9597	1.0000

Eigenvalues of the Correlation Matrix				
	Eigenvalue	Difference	Proportion	Cumulative
1	2.90808114	2.82438377	0.9694	0.9694
2	0.08369737	0.07547588	0.0279	0.9973
3	0.00822149		0.0027	1.0000

Eigenvectors			
	Prin1	Prin2	Prin3
knowhow	0.576251	-.618121	0.534660
problem_solving	0.584343	-.145758	-.798310
accountability	0.571383	0.772451	0.277201

The PRINCOMP Procedure



The PRINCOMP Procedure

Observations	67
Variables	3

Simple Statistics

	knowhow	problem_solving	accountability
Mean	0.000000000	0.000000000	0.000000000
StD	1.000000000	1.000000000	1.000000000

Correlation Matrix

	knowhow	problem_solving	accountability
knowhow	1.0000	0.9833	0.9188
problem_solving	0.9833	1.0000	0.9597
accountability	0.9188	0.9597	1.0000

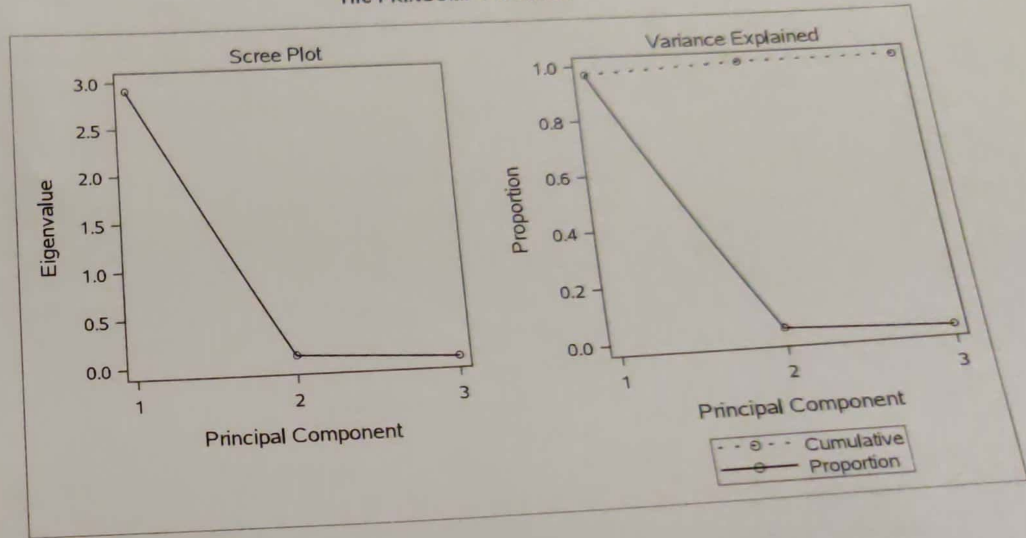
Eigenvalues of the Correlation Matrix

	Eigenvalue	Difference	Proportion	Cumulative
1	2.90808114	2.82438377	0.9694	0.9694
2	0.08369737	0.07547588	0.0279	0.9973
3	0.00822149		0.0027	1.0000

Eigenvectors

	Prin1	Prin2	Prin3
knowhow	0.576251	-.618121	0.534660
problem_solving	0.584343	-.145758	-.798310
accountability	0.571383	0.772451	0.277201

The PRINCOMP Procedure



The REG Procedure
Model: MODEL1
Dependent Variable: Prin1

Number of Observations Read	67
Number of Observations Used	67

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

Analysis of Variance

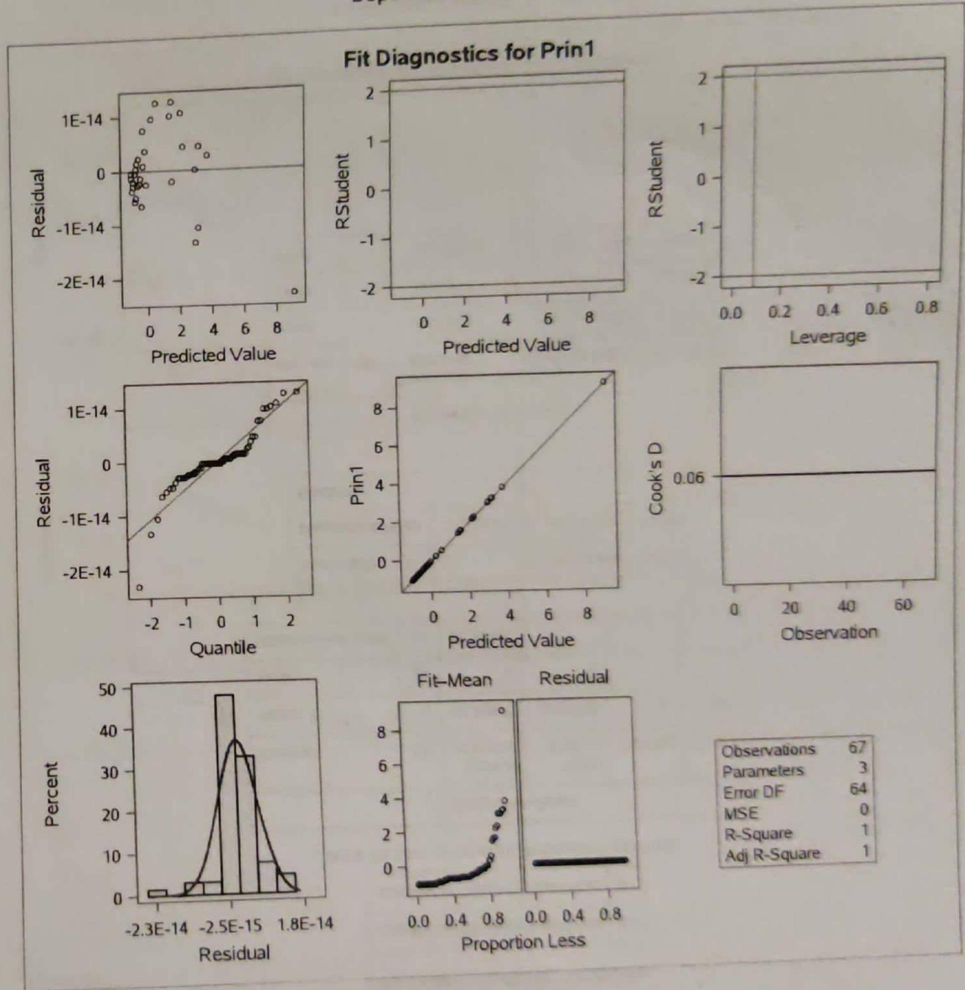
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	191.93336	63.97779	Infty	<.0001
Error	64	0	0		
Uncorrected Total	67	191.93336			

Root MSE	0	R-Square	1.0000
Dependent Mean	3.24782E-16	Adj R-Sq	1.0000
Coeff Var	0		

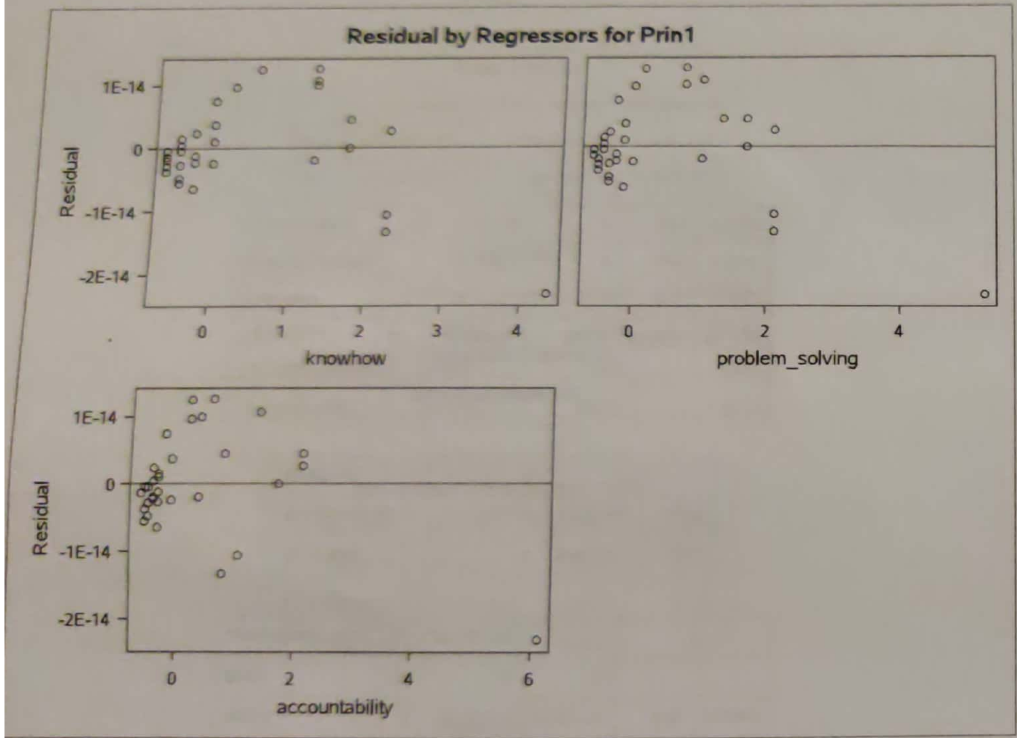
Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
knowhow	1	0.57625	0	Infty	<.0001
problem_solving	1	0.58434	0	Infty	<.0001
accountability	1	0.57138	0	Infty	<.0001

The REG Procedure
Model: MODEL1
Dependent Variable: Prin1



The REG Procedure
Model: MODEL1
Dependent Variable: Prin1



The REG Procedure
Model: MODEL1
Dependent Variable: knowhow

Number of Observations Read	67
Number of Observations Used	67

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

Analysis of Variance

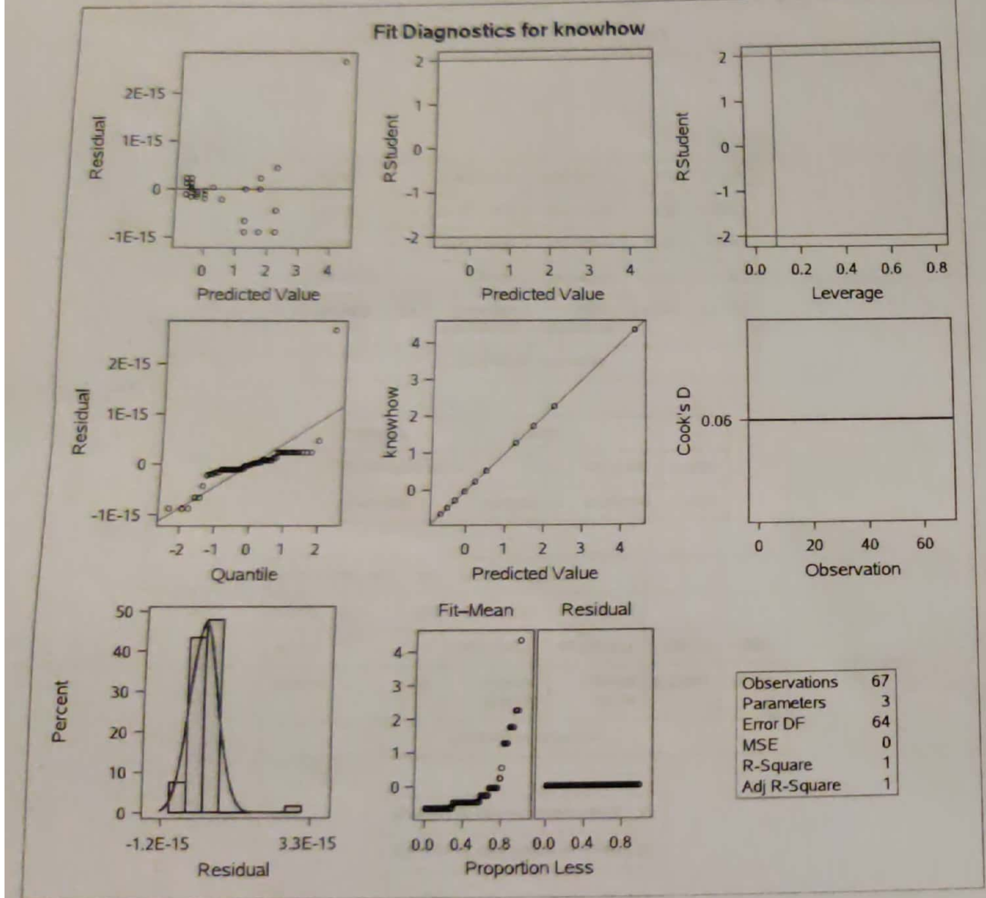
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	66.00000	22.00000	Infty	<.0001
Error	64	0	0		
Uncorrected Total	67	66.00000			

Root MSE	0	R-Square	1.0000
Dependent Mean	3.67865E-16	Adj R-Sq	1.0000
Coeff Var	0		

Parameter Estimates

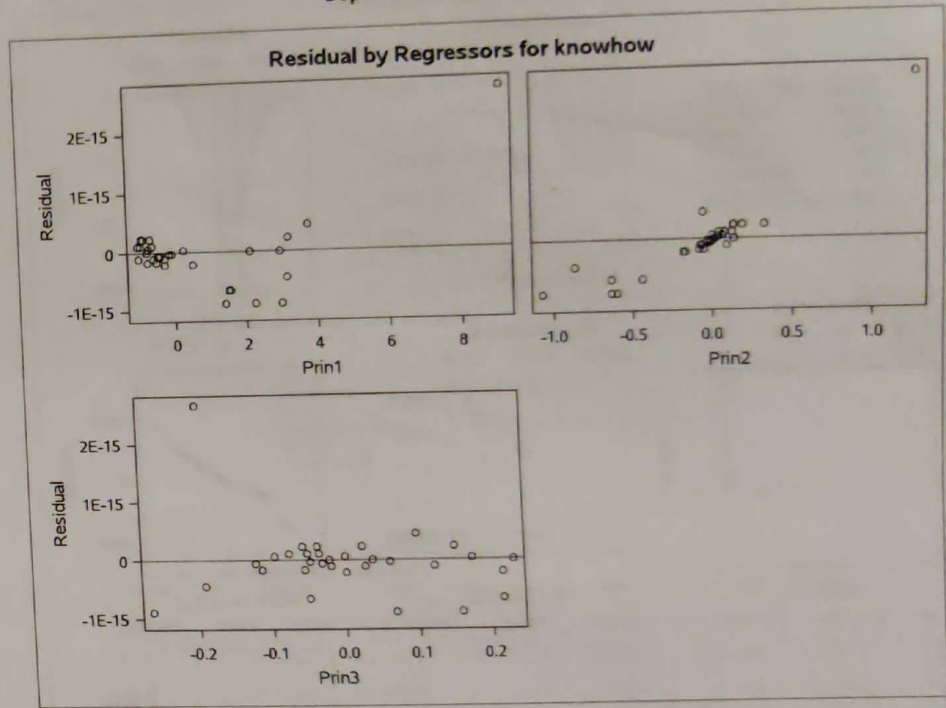
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Prin1	1	0.57625	0	Infty	<.0001
Prin2	1	-0.61812	0	-Infty	<.0001
Prin3	1	0.53466	0	Infty	<.0001

The REG Procedure
Model: MODEL1
Dependent Variable: knowhow



Observations	67
Parameters	3
Error DF	64
MSE	0
R-Square	1
Adj R-Square	1

The REG Procedure
Model: MODEL1
Dependent Variable: knowhow



The REG Procedure
Model: MODEL1
Dependent Variable: salary

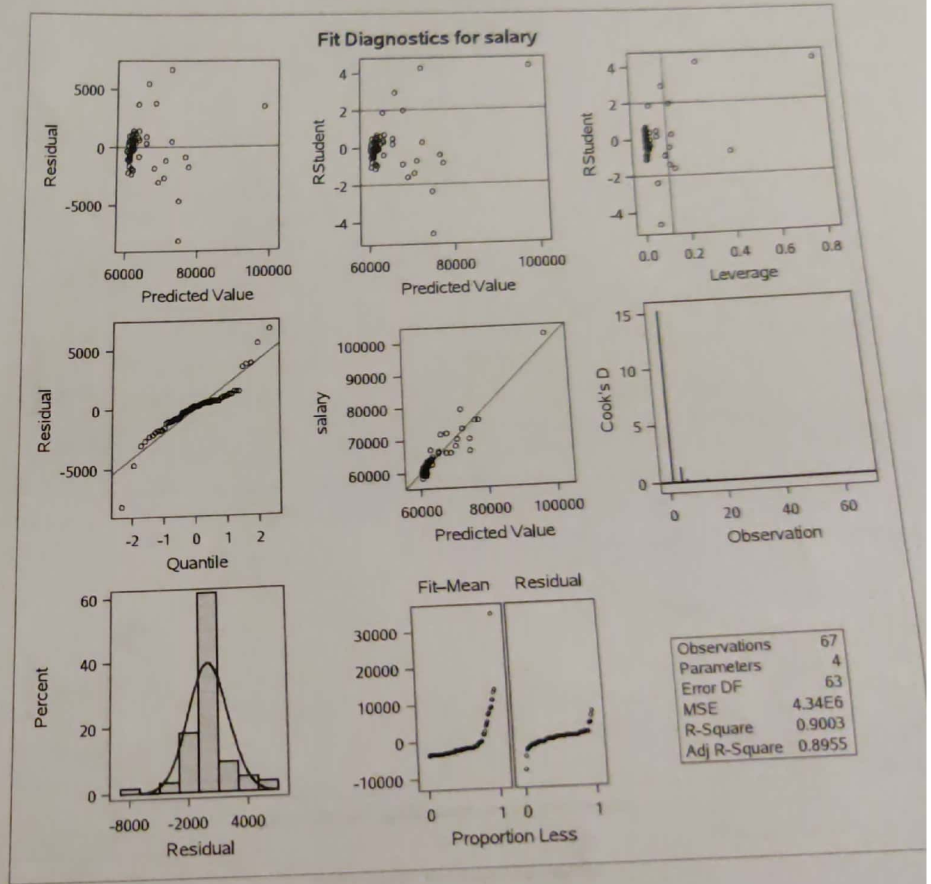
Number of Observations Read	67
Number of Observations Used	67

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	2465105931	821701977	189.55	<.0001
Error	63	273111655	4335106		
Corrected Total	66	2738217587			

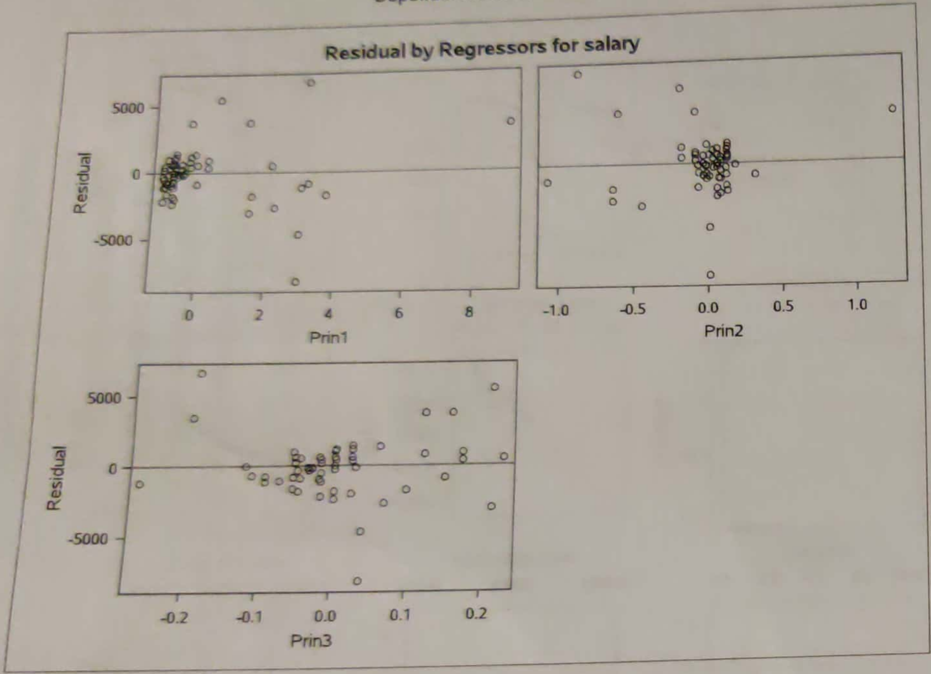
Root MSE	2082.09165	R-Square	0.9003
Dependent Mean	63929	Adj R-Sq	0.8955
Coeff Var	3.25686		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	63929	254.36798	251.33	<.0001
Prin1	1	3557.20641	150.28811	23.67	<.0001
Prin2	1	2316.12408	885.87403	2.61	0.0112
Prin3	1	3540.61136	2826.52316	1.25	0.2150

The REG Procedure
Model: MODEL1
Dependent Variable: salary



The REG Procedure
Model: MODEL1
Dependent Variable: salary



95/100

The REG Procedure
Model: MODEL1
Dependent Variable: salary

Number of Observations Read	67
Number of Observations Used	67

Analysis of Variance

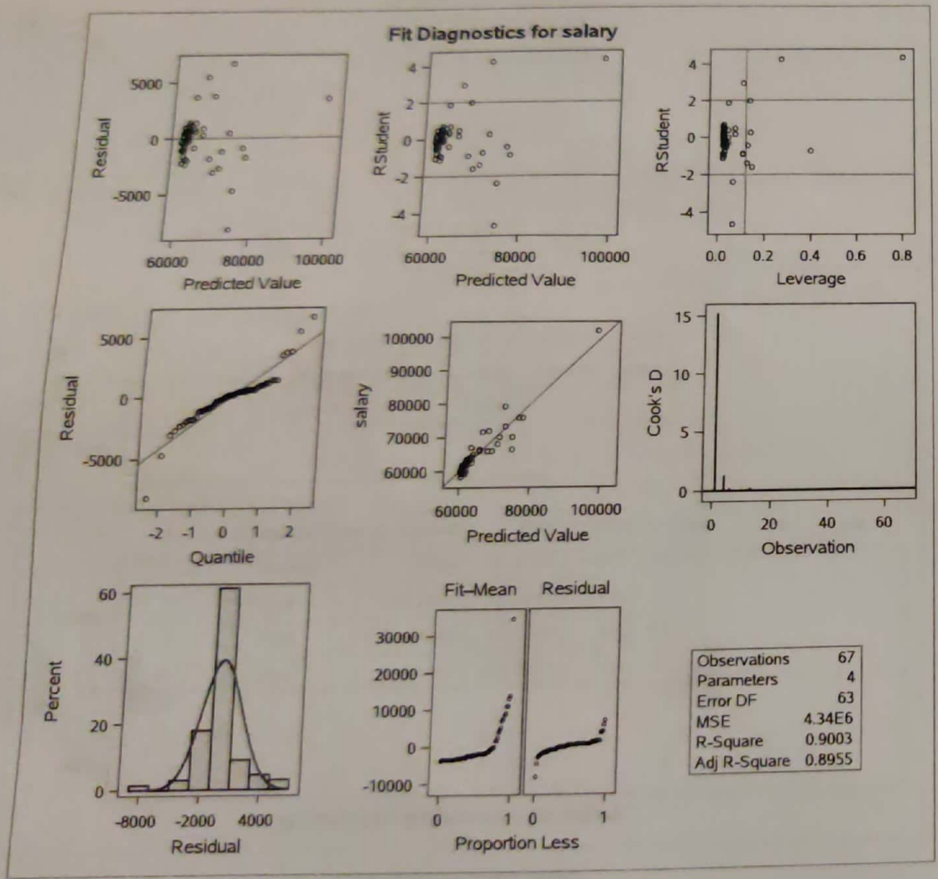
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	2465105931	821701977	189.55	<.0001
Error	63	273111655	4335106		
Corrected Total	66	2738217587			

Root MSE	2082.09165	R-Square	0.9003
Dependent Mean	63929	Adj R-Sq	0.8955
Coeff Var	3.25686		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Type I SS	Type II SS	Variance Inflation
Intercept	1	63929	254.36798	251.33	<.0001	2.738263E11	2.738263E11	0
Prin1	1	3557.20641	150.28811	23.67	<.0001	2428670447	2428670447	1.00000
Prin2	1	2316.12408	885.87403	2.61	0.0112	29633257	29633257	1.00000
Prin3	1	3540.61136	2826.52316	1.25	0.2150	6802227	6802227	1.00000

The REG Procedure
Model: MODEL1
Dependent Variable: salary



The REG Procedure
Model: MODEL1
Dependent Variable: salary