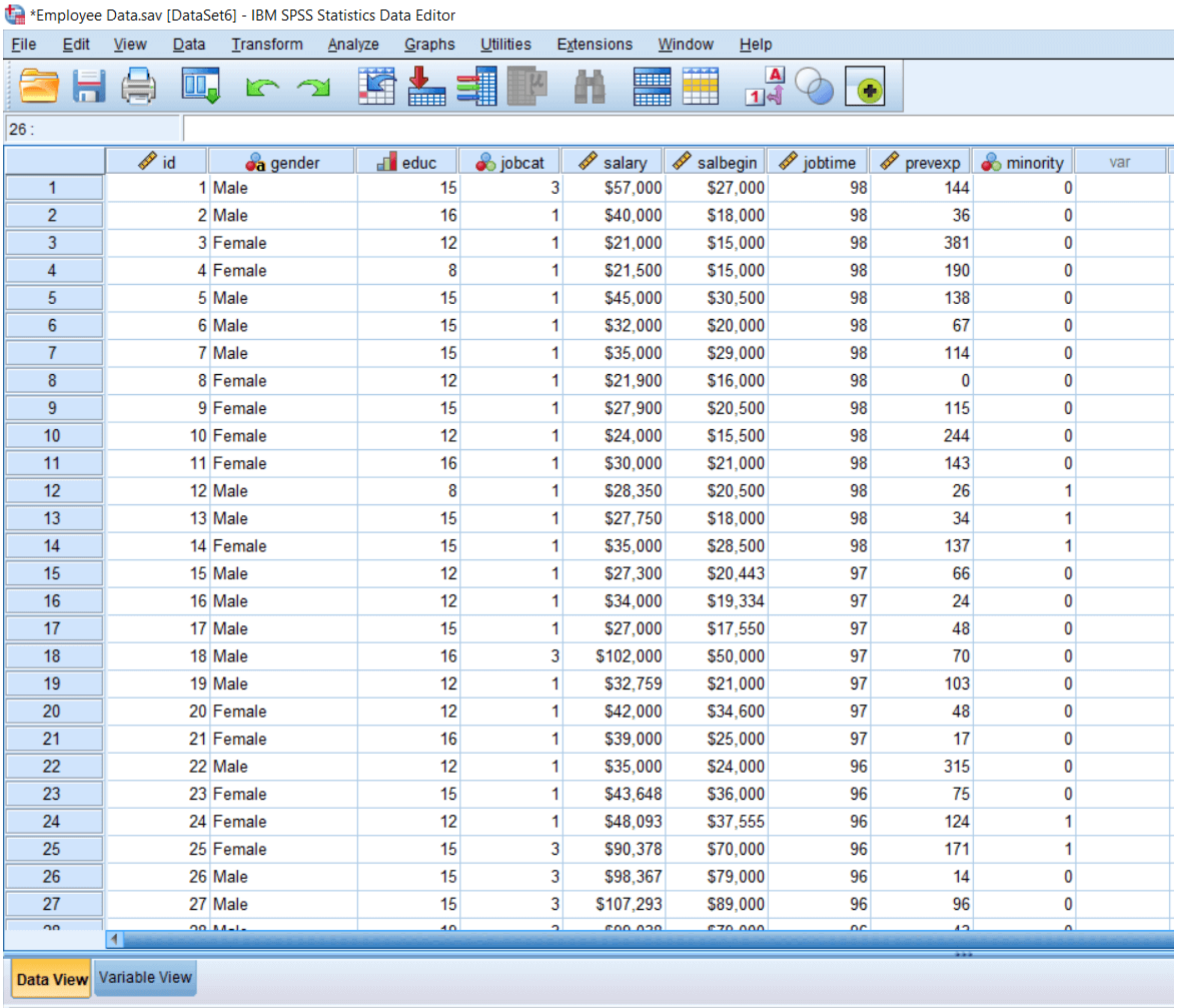
One-Way ANOVA

One-Way ANOVA in SPSS

- In this section, we will learn how to conduct **One-Way ANOVA** in SPSS. The theory of One-Way ANOVA was given by British Mathematician or statistician **Ronald Fisher**. He gives One-Way ANOVA to identify the significant differences between more than two groups. Whenever we have more than two groups, we cannot apply a **t-test**. So, in that case, we have to go to **One-Way ANOVA**. Generally, more than **two groups** mean more than **two levels** of one independent variable.
- **For example**, suppose there is a **researcher** who wants to determine the effect of three different kinds of **therapies** or **medicines** on the patient's condition, such as depression. Suppose he treats the patient for **depression** and has three or more than **three therapies**. In that case, we want to determine whether these three different therapies lead to significantly different results. We will check whether **one therapy** leads to significantly more **improvement** in the patient's condition as compared to other **therapy**. In this case, we have just **two variables**. The first variable is a **therapy** that is our **independent** variable, and the second variable is **depression,** that is our **dependent**
- **Therapy** has three levels, which are **Psychodynamic** therapy, **behavioral** therapy, and **cognitive** So we have **three levels** of a single variable. In that case, we want to determine whether these three different levels have a significantly different influence on the dependent variable.

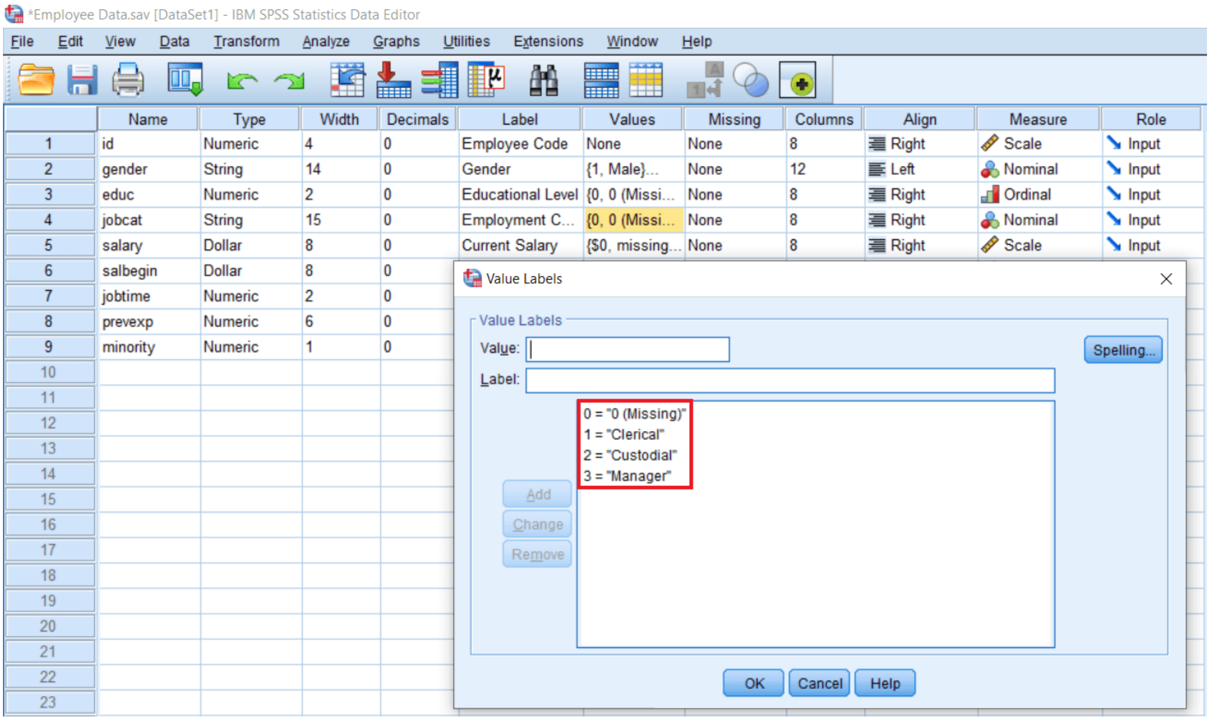
Calculating One-Way ANOVA in SPSS

To calculate One-Way ANOVA, we will take the **Employee data set,** which is given below:

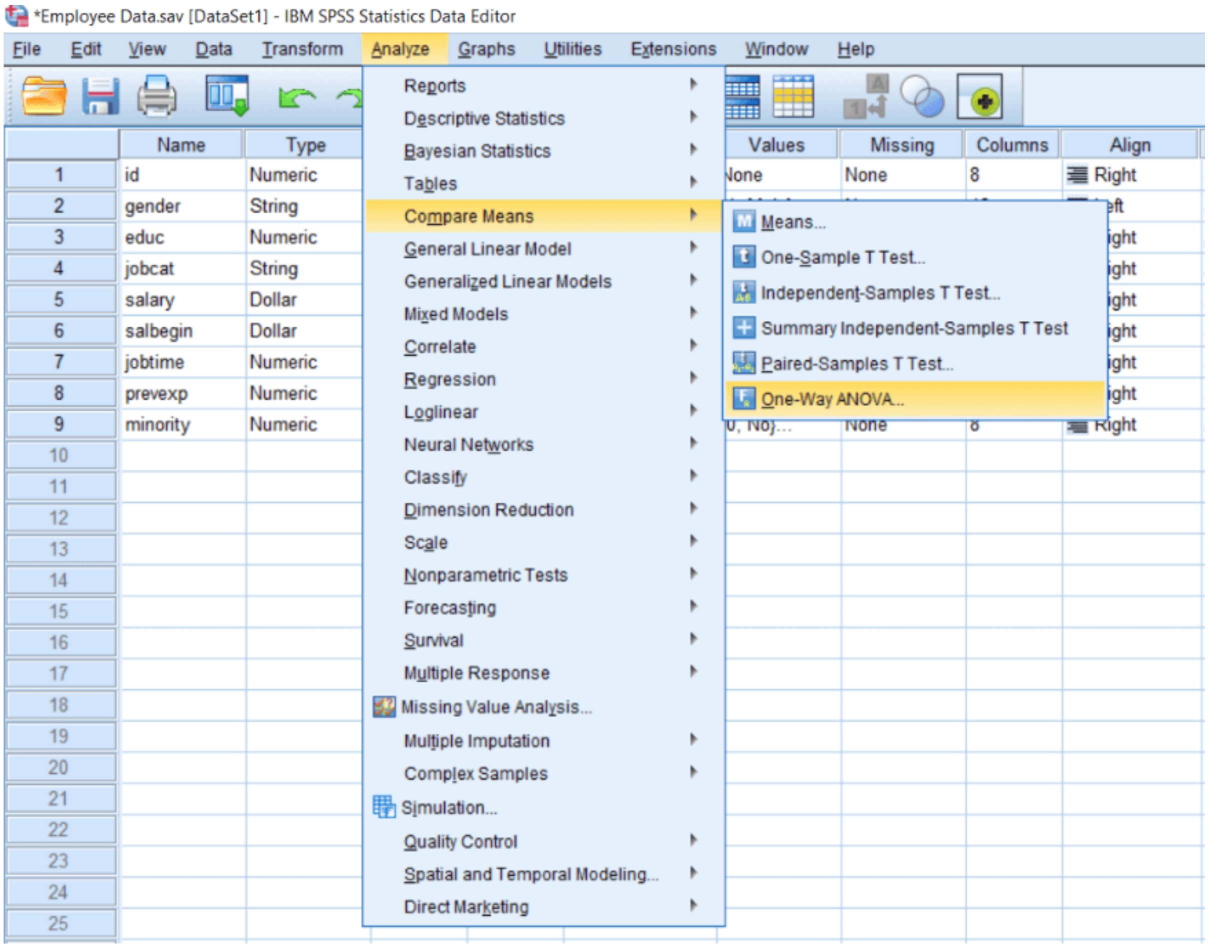


	id	gender	educ	jobcat	salary	salbegin	jobtime	prevexp	minority	var
1	1	Male	15	3	$57,000	$27,000	98	144	0	
2	2	Male	16	1	$40,000	$18,000	98	36	0	
3	3	Female	12	1	$21,000	$15,000	98	381	0	
4	4	Female	8	1	$21,500	$15,000	98	190	0	
5	5	Male	15	1	$45,000	$30,500	98	138	0	
6	6	Male	15	1	$32,000	$20,000	98	67	0	
7	7	Male	15	1	$35,000	$29,000	98	114	0	
8	8	Female	12	1	$21,900	$16,000	98	0	0	
9	9	Female	15	1	$27,900	$20,500	98	115	0	
10	10	Female	12	1	$24,000	$15,500	98	244	0	
11	11	Female	16	1	$30,000	$21,000	98	143	0	
12	12	Male	8	1	$28,350	$20,500	98	26	1	
13	13	Male	15	1	$27,750	$18,000	98	34	1	
14	14	Female	15	1	$35,000	$28,500	98	137	1	
15	15	Male	12	1	$27,300	$20,443	97	66	0	
16	16	Male	12	1	$34,000	$19,334	97	24	0	
17	17	Male	15	1	$27,000	$17,550	97	48	0	
18	18	Male	16	3	$102,000	$50,000	97	70	0	
19	19	Male	12	1	$32,759	$21,000	97	103	0	
20	20	Female	12	1	$42,000	$34,600	97	48	0	
21	21	Female	16	1	$39,000	$25,000	97	17	0	
22	22	Male	12	1	$35,000	$24,000	96	315	0	
23	23	Female	15	1	$43,648	$36,000	96	75	0	
24	24	Female	12	1	$48,093	$37,555	96	124	1	
25	25	Female	15	3	$90,378	$70,000	96	171	1	
26	26	Male	15	3	$98,367	$79,000	96	14	0	
27	27	Male	15	3	$107,293	$89,000	96	96	0	

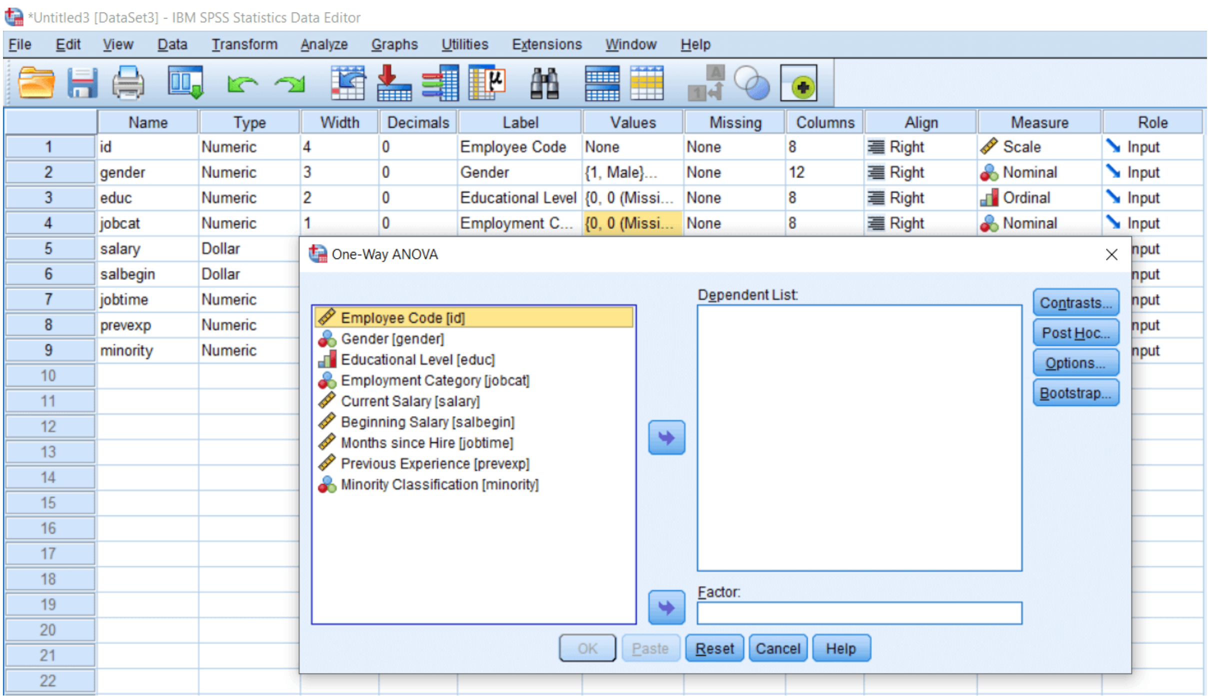
In this data set, we have a variable called **jobcat**. Now, if we look at the **jobcat variable**, we find that jobcat has **three levels**. If we click on the Values option, there are people from three different job categories, i.e., **Clerical** level, **Custodial** level, and **Managerial** level. **0** stands for the **missing value**. That is a good way of defining missing value. 0 has been defined as a missing value, and they have kept in the missing value option. It means if we calculate any **stat,** for example, we want to do **One-Way ANOVA**, 0 will not be counted as a variable.



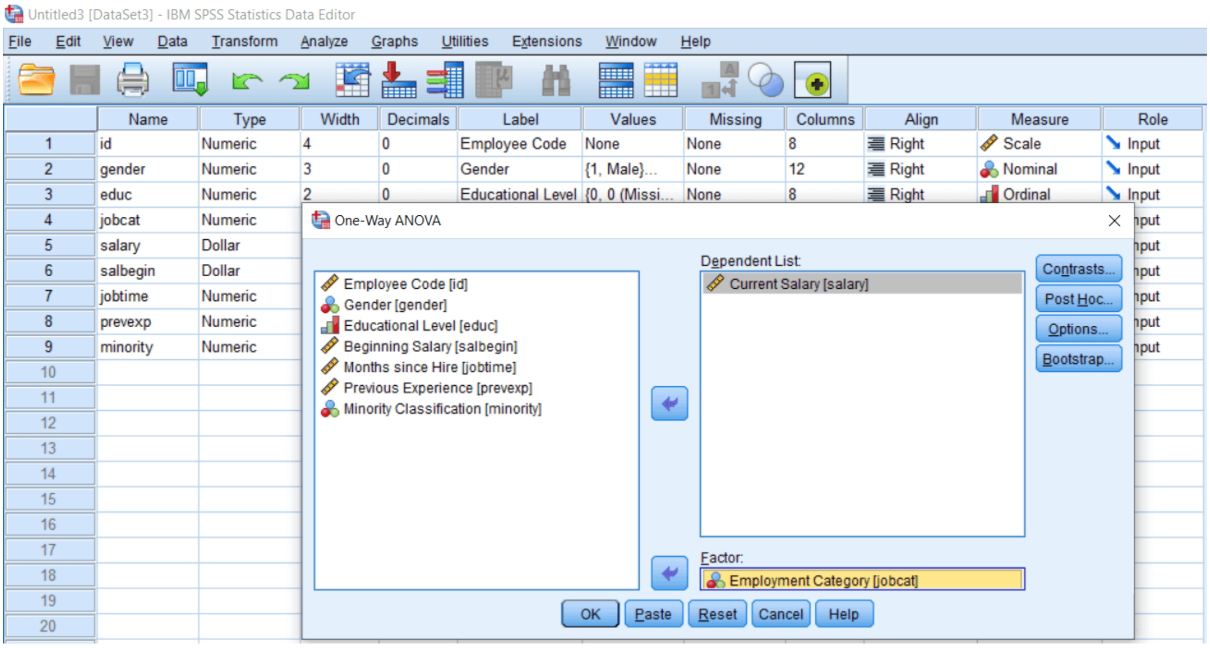
We want to find out that people belonging to different job categories will be drawing significantly **different** amounts of **salary**. To be more precise, we want to make a **test**. In which we can say that **Managers** are drawing significantly **more** amount to **salary** as compare to **custodial** or **clerical** people. So we hypothesize that when we move from the **clerical** level to the **managerial** level, there is an increase in salary. We can test it by using **One-Way ANOVA**. So we have **two variables** here. One is an **independent** variable that is **jobcat**, which has **three levels** and one **dependent** variable, i.e., **current salary**. Now once all variables are defined, we will move to the calculation of **One-Way ANOVA**. To calculate **One-Way ANOVA,** we will go **Analyze** menu, and **Compare Means** and then **One-Way ANOVA**.



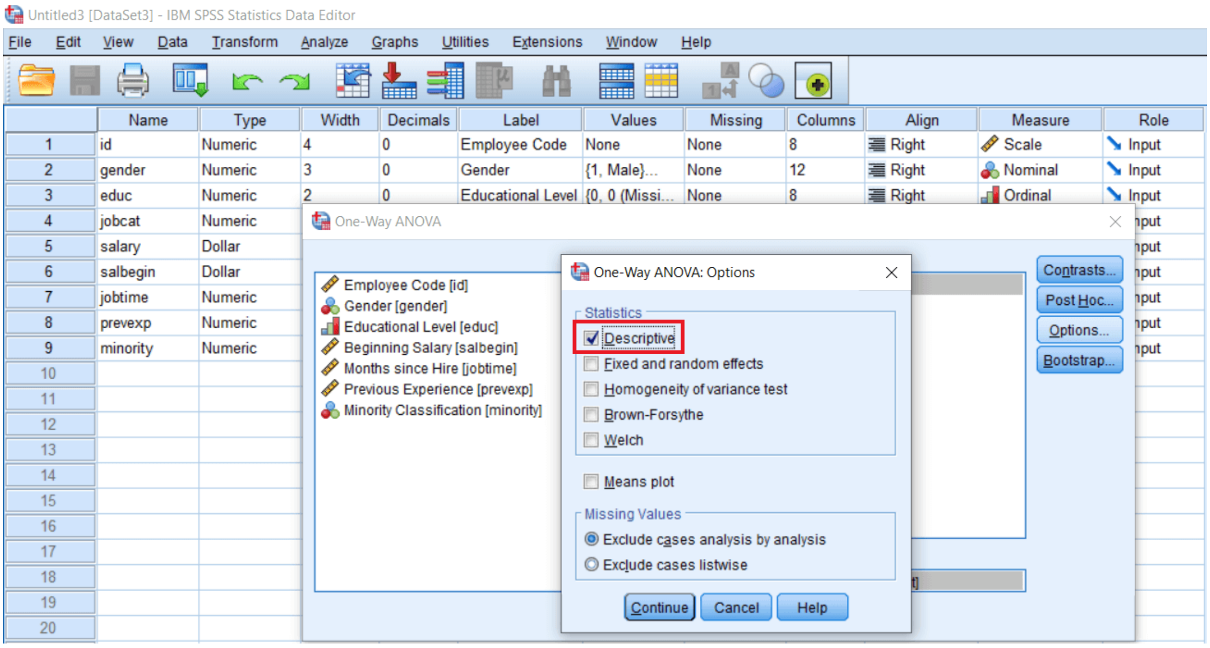
The symbol of **One-Way ANOVA** is **F a**. So **F** stands for the **Ronald fisher**, the person who gives this test. When we click on **One-Way ANOVA**, we will see a dialog box like this:



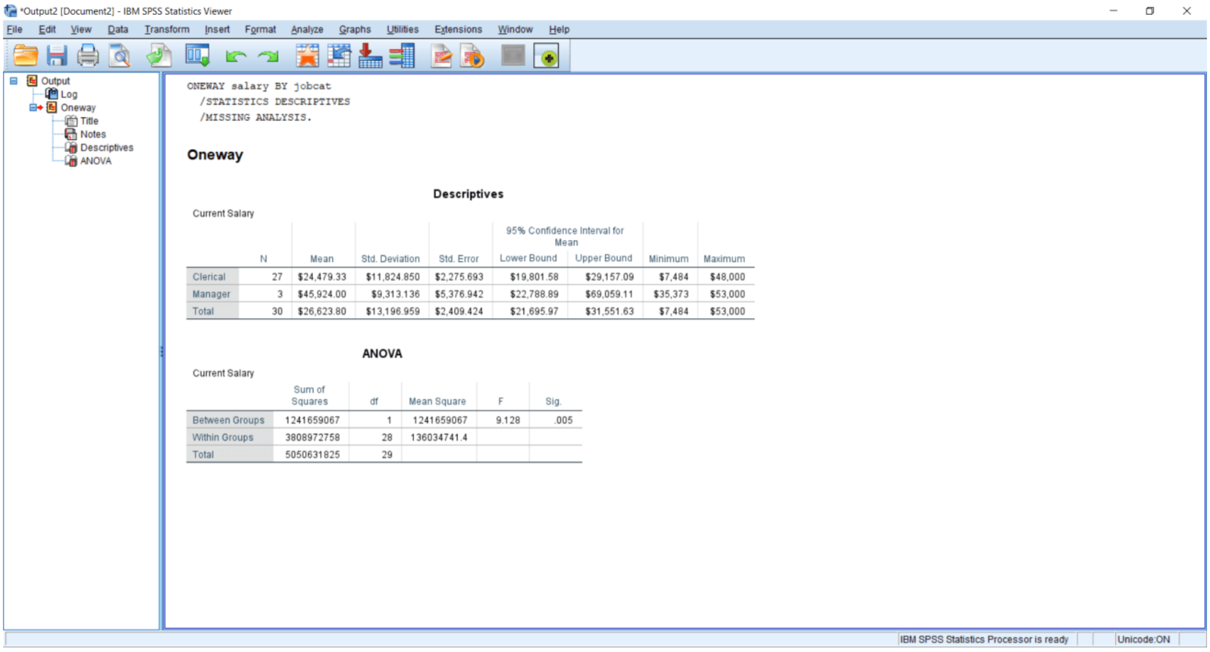
All the **variables** are populated on the **left-hand side**. So we have to select our **dependent** variable in the **Dependent List** and **independent** variable in **Factor**. So our **dependent** variable is the **Current salary,** and the **independent** variable is **Job category,** so we will select it.



By going to the **Options tab**, we can look at the **Descriptive statistics** like this:



Now we are ready to test the significant differences between three **salary groups**, three **job categories**. So click on, **Ok**. The following is the **output** for **One-Way ANOVA**.



*Output2 [Document2] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Output
Log
Oneway
Title
Notes
Descriptives
ANOVA

```
ONEWAY salary BY jobcat
  /STATISTICS DESCRIPTIVES
  /MISSING ANALYSIS.
```

Oneway

Descriptives

Current Salary

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Clerical	27	$24,479.33	$11,824.850	$2,275.693	$19,801.58	$29,157.09	$7,484	$48,000
Manager	3	$45,924.00	$9,313.136	$5,376.942	$22,788.89	$69,059.11	$35,373	$53,000
Total	30	$26,623.80	$13,196.959	$2,409.424	$21,695.97	$31,551.63	$7,484	$53,000

ANOVA

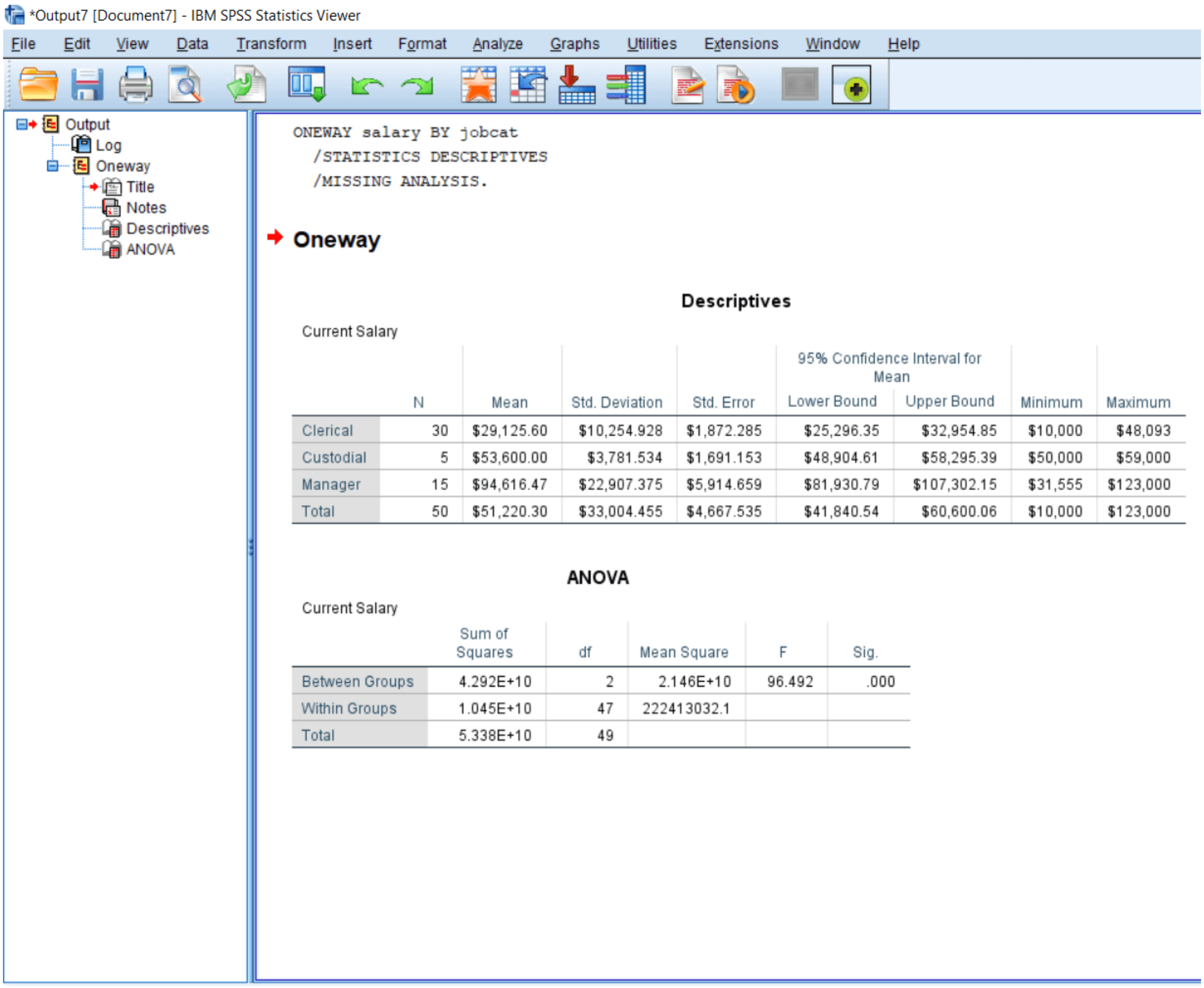
Current Salary

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1241659067	1	1241659067	9.128	.005
Within Groups	3808972758	28	136034741.4		
Total	5050631825	29			

IBM SPSS Statistics Processor is ready | Unicode:ON

Output of One-Way ANOVA

In this section, we will learn the output of **One-Way ANOVA**. The output of **One-Way ANOVA** is given below:




```
ONEWAY salary BY jobcat
  /STATISTICS DESCRIPTIVES
  /MISSING ANALYSIS.
```

Oneway

Descriptives

Current Salary

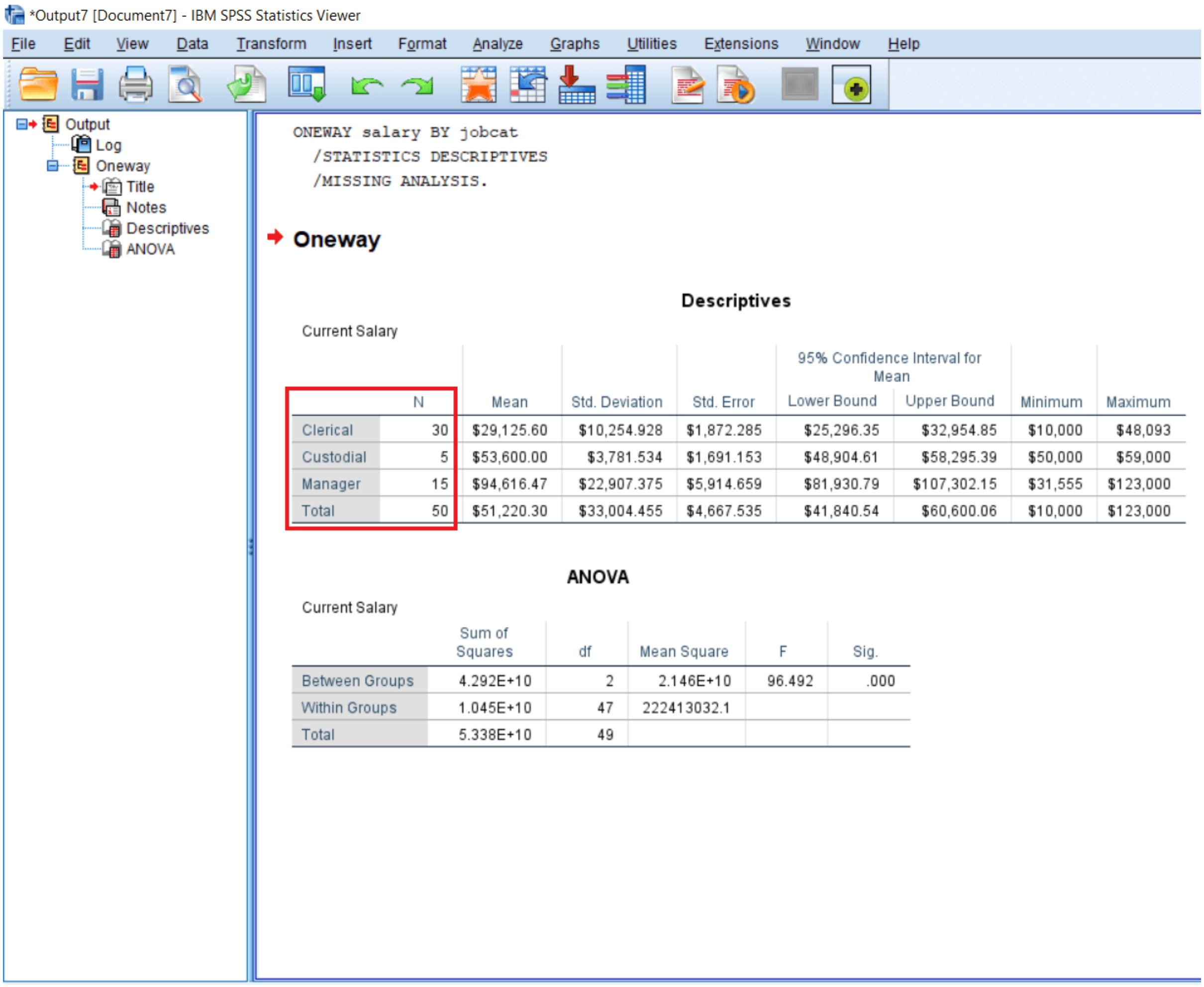
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
Clerical	30	$29,125.60	$10,254.928	$1,872.285	$25,296.35	$32,954.85	$10,000	$48,093
Custodial	5	$53,600.00	$3,781.534	$1,691.153	$48,904.61	$58,295.39	$50,000	$59,000
Manager	15	$94,616.47	$22,907.375	$5,914.659	$81,930.79	$107,302.15	$31,555	$123,000
Total	50	$51,220.30	$33,004.455	$4,667.535	$41,840.54	$60,600.06	$10,000	$123,000

ANOVA

Current Salary

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4.292E+10	2	2.146E+10	96.492	.000
Within Groups	1.045E+10	47	222413032.1		
Total	5.338E+10	49			

In the above output, we can see only three **job categories** are appearing, i.e., **Clerical, Custodial**, and **Manager**. The missing value **0** is not available in the output. **30** people belong to job **category 1,** i.e., **Clerical, 5** belonging to job **category 2,** i.e., **Custodial**, and **15** belonging to job **category 3,** i.e., **Managers**. We are having a slightly lesser number of people in job category 2.




```
ONEWAY salary BY jobcat
  /STATISTICS DESCRIPTIVES
  /MISSING ANALYSIS.
```

Oneway

Descriptives

Current Salary

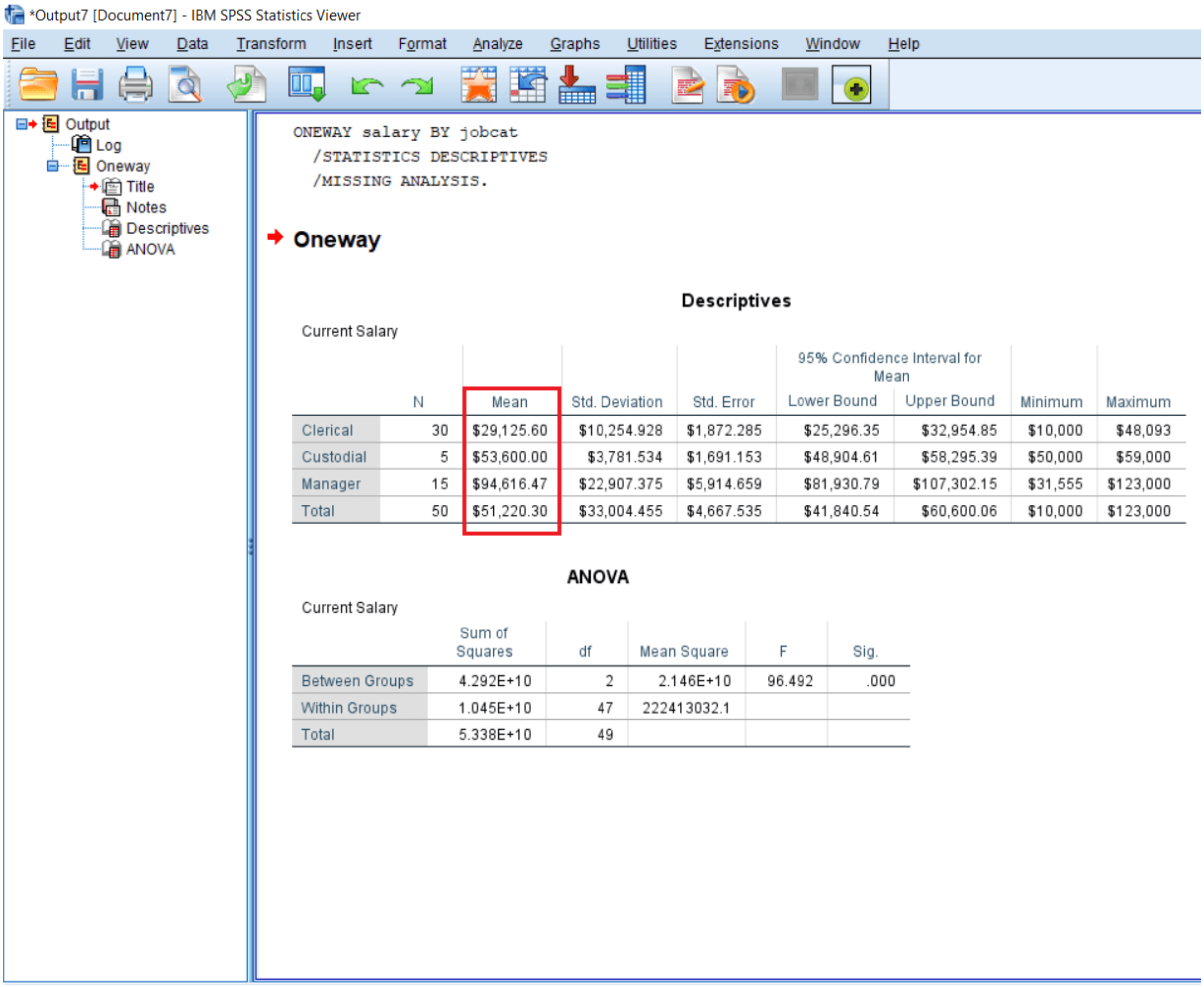
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	Upper Bound	Minimum	Maximum
Clerical	30	$29,125.60	$10,254.928	$1,872.285	$25,296.35	$32,954.85	$10,000	$48,093
Custodial	5	$53,600.00	$3,781.534	$1,691.153	$48,904.61	$58,295.39	$50,000	$59,000
Manager	15	$94,616.47	$22,907.375	$5,914.659	$81,930.79	$107,302.15	$31,555	$123,000
Total	50	$51,220.30	$33,004.455	$4,667.535	$41,840.54	$60,600.06	$10,000	$123,000

ANOVA

Current Salary

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4.292E+10	2	2.146E+10	96.492	.000
Within Groups	1.045E+10	47	222413032.1		
Total	5.338E+10	49			

Now we will see it has any recursion for our analysis. **Average salaries** are **29,125** dollars, **53,600** dollars, and **94,616** dollars. If we look at these **Means scores**, we will see the impression that there are differences between people's salaries from three different job categories. But if we move from **Clerks** to **Custodial** employees, the amount of difference is not huge. But if we move from the **Custodial** employees to the **Manager**, there is a huge significant difference, more than double the salary difference. So, we are expecting a significant difference between the groups. So that's how we can guess about the significant differences just by looking at the **Mean scores**.


*Output7 [Document7] - IBM SPSS Statistics Viewer

File Edit View Data Transform Insert Format Analyze Graphs Utilities Extensions Window Help

Output
Log
Oneway
Title
Notes
Descriptives
ANOVA

```
ONEWAY salary BY jobcat
  /STATISTICS DESCRIPTIVES
  /MISSING ANALYSIS.
```

Oneway

Descriptives

Current Salary

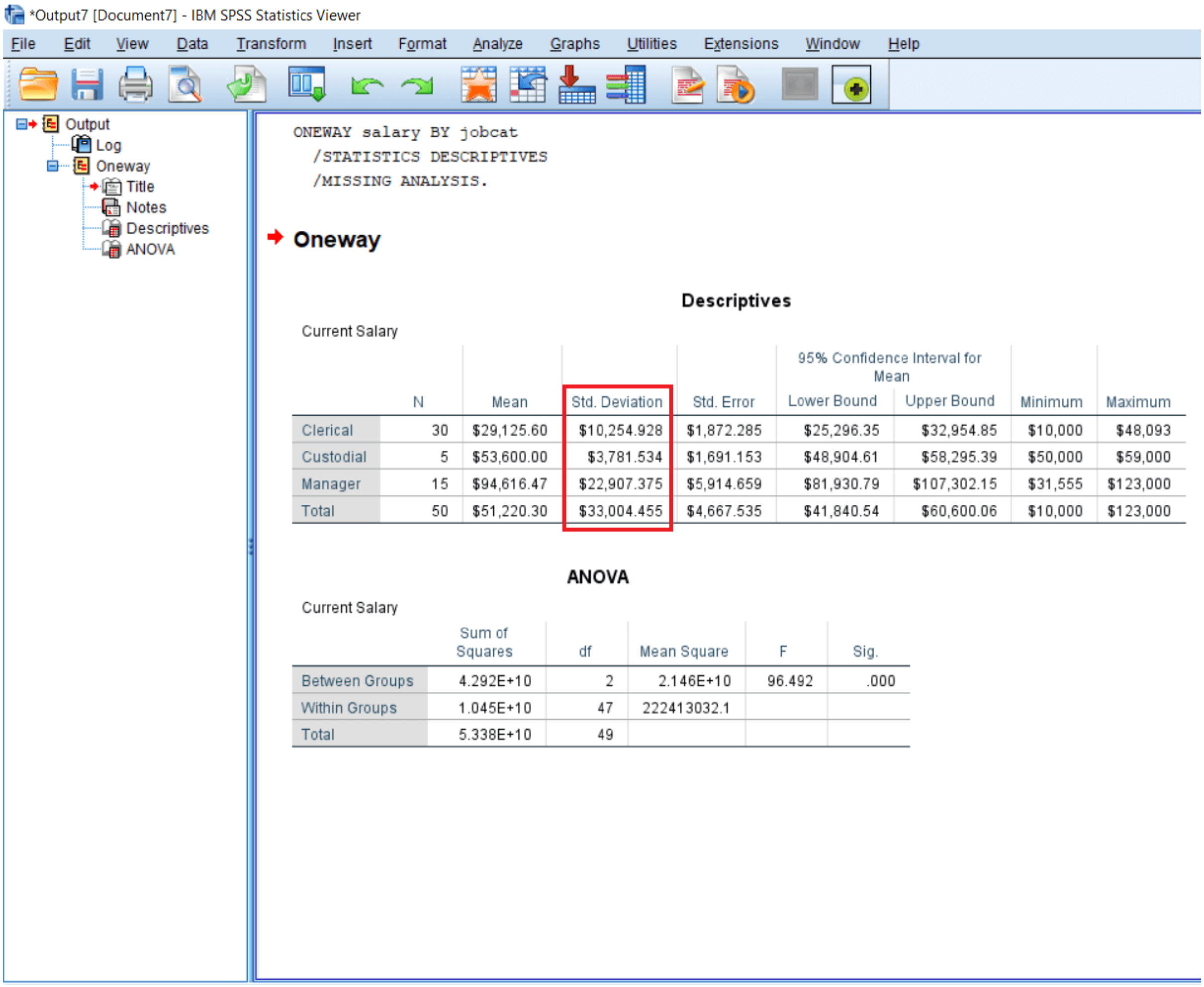
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
Clerical	30	$29,125.60	$10,254.928	$1,872.285	$25,296.35	$32,954.85	$10,000	$48,093
Custodial	5	$53,600.00	$3,781.534	$1,691.153	$48,904.61	$58,295.39	$50,000	$59,000
Manager	15	$94,616.47	$22,907.375	$5,914.659	$81,930.79	$107,302.15	$31,555	$123,000
Total	50	$51,220.30	$33,004.455	$4,667.535	$41,840.54	$60,600.06	$10,000	$123,000

ANOVA

Current Salary

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4.292E+10	2	2.146E+10	96.492	.000
Within Groups	1.045E+10	47	222413032.1		
Total	5.338E+10	49			

The **Standard deviation** in the case of **Managers** is quite **high** compared to the rest of two groups. So, managers are drawing on an average more salary, but there is a huge variation in the managerial category compared to other groups.




```
ONEWAY salary BY jobcat
  /STATISTICS DESCRIPTIVES
  /MISSING ANALYSIS.
```

Oneway

Descriptives

Current Salary

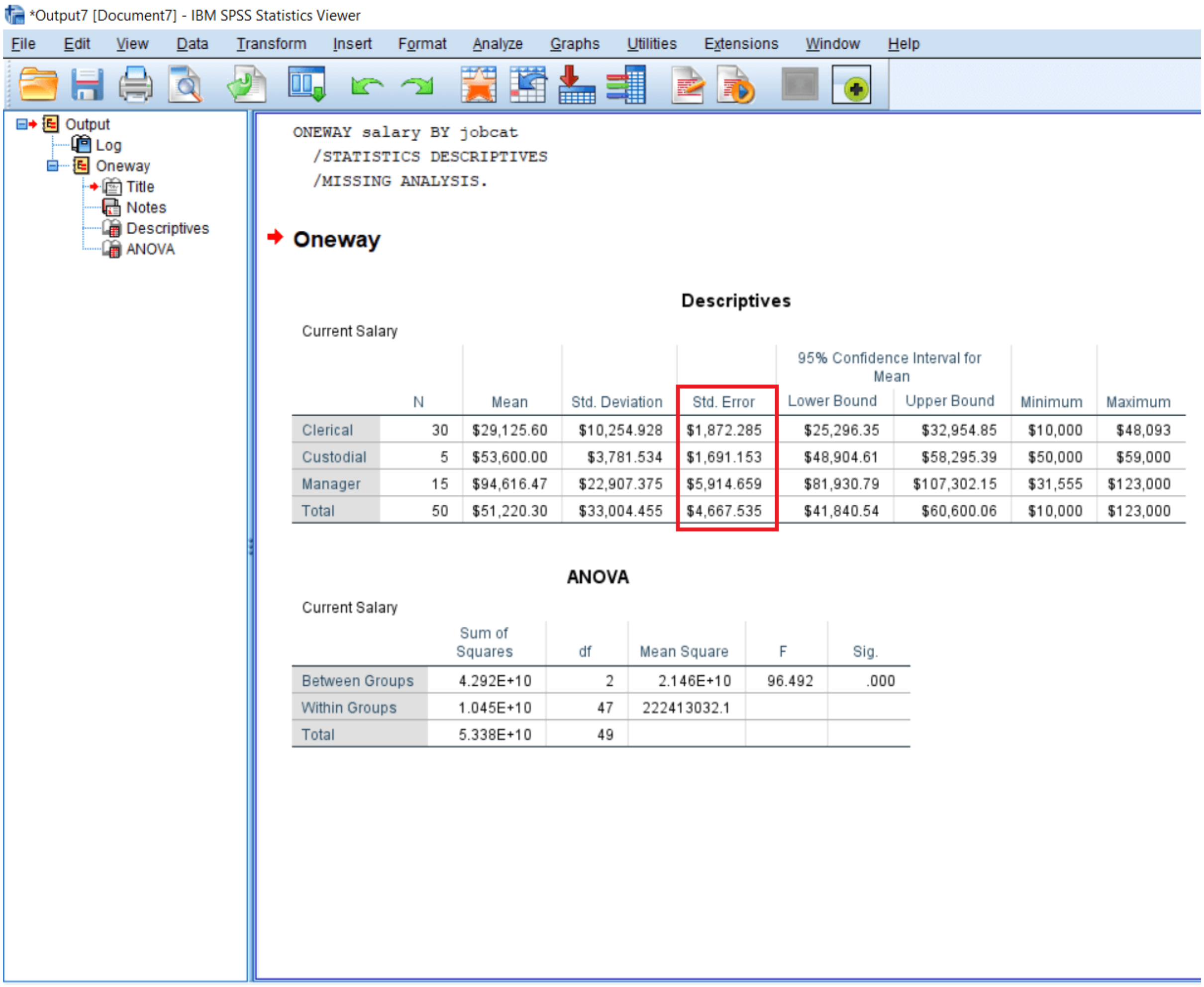
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean Lower Bound	95% Confidence Interval for Mean Upper Bound	Minimum	Maximum
Clerical	30	$29,125.60	$10,254.928	$1,872.285	$25,296.35	$32,954.85	$10,000	$48,093
Custodial	5	$53,600.00	$3,781.534	$1,691.153	$48,904.61	$58,295.39	$50,000	$59,000
Manager	15	$94,616.47	$22,907.375	$5,914.659	$81,930.79	$107,302.15	$31,555	$123,000
Total	50	$51,220.30	$33,004.455	$4,667.535	$41,840.54	$60,600.06	$10,000	$123,000

ANOVA

Current Salary

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4.292E+10	2	2.146E+10	96.492	.000
Within Groups	1.045E+10	47	222413032.1		
Total	5.338E+10	49			

The **Standard error** is shown in the following image. **Standard error** refers to the standard deviation of the sampling distribution of mean. So, it's an indication of the amount of error in measurement. So, the smaller it is better for us.




```
ONEWAY salary BY jobcat
  /STATISTICS DESCRIPTIVES
  /MISSING ANALYSIS.
```

Oneway

Descriptives

Current Salary

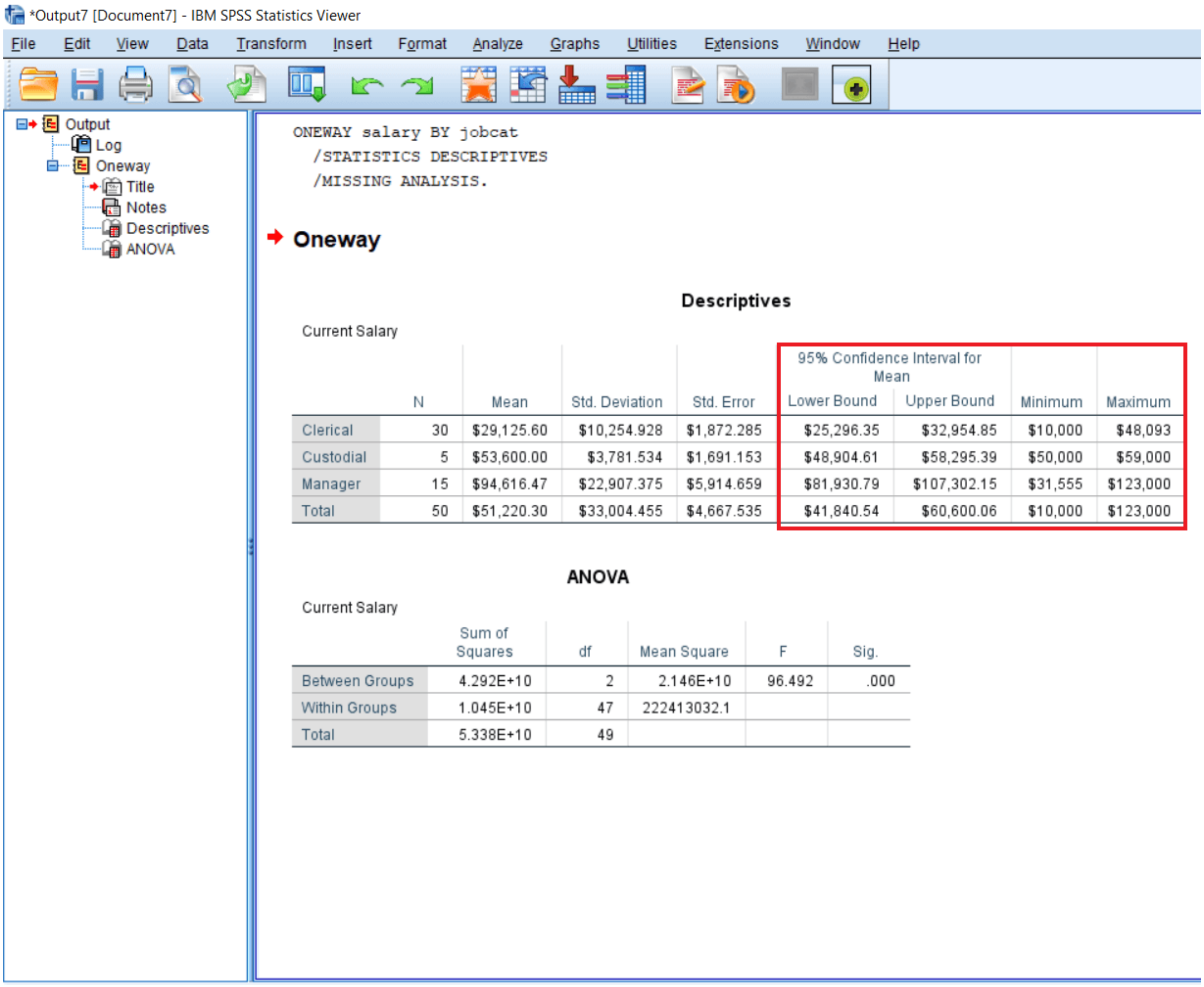
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Clerical	30	$29,125.60	$10,254.928	$1,872.285	$25,296.35	$32,954.85	$10,000	$48,093
Custodial	5	$53,600.00	$3,781.534	$1,691.153	$48,904.61	$58,295.39	$50,000	$59,000
Manager	15	$94,616.47	$22,907.375	$5,914.659	$81,930.79	$107,302.15	$31,555	$123,000
Total	50	$51,220.30	$33,004.455	$4,667.535	$41,840.54	$60,600.06	$10,000	$123,000

ANOVA

Current Salary

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4.292E+10	2	2.146E+10	96.492	.000
Within Groups	1.045E+10	47	222413032.1		
Total	5.338E+10	49			

There is **95% confidence information**. So, we can see none of the confidence information is having 0. So, we don't have a positive or negative value on either side, and these are the minimum and maximum amount of salary drawn by different groups. So, in the case of **Manager,** the difference is very high, **31 thousand** to **123 thousand** that's the descriptive scenario of output.




```
ONEWAY salary BY jobcat
  /STATISTICS DESCRIPTIVES
  /MISSING ANALYSIS.
```

Oneway

Descriptives

Current Salary

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Clerical	30	$29,125.60	$10,254.928	$1,872.285	$25,296.35	$32,954.85	$10,000	$48,093
Custodial	5	$53,600.00	$3,781.534	$1,691.153	$48,904.61	$58,295.39	$50,000	$59,000
Manager	15	$94,616.47	$22,907.375	$5,914.659	$81,930.79	$107,302.15	$31,555	$123,000
Total	50	$51,220.30	$33,004.455	$4,667.535	$41,840.54	$60,600.06	$10,000	$123,000

ANOVA

Current Salary

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	4.292E+10	2	2.146E+10	96.492	.000
Within Groups	1.045E+10	47	222413032.1		
Total	5.338E+10	49			