T.C.
Yeditepe University

Graduate School of Natural and Applied Sciences

Computer Engineering Department

The Object Project

SECTION-1

Batuhan Düzgün

School Number: 20163505004

2017-2018

1. Aim of the Research

The aim of research is that research to examine what changes prices of Desktop Central Processing Unit (CPU).

2. Target Population Definition

In this research, the target population is the Desktop Central Processing Unit (CPU) products. The up to date price of Desktop CPUs will be taken into account.

3. Core Concept Definition

Core concept in this research is to research to investigate what changes the up to date price of Desktop Central Processing Unit.

4. Definition of External Variables in relation to core concept

External Variables are CPU core count, Socket type, Cache Memory Size, CPU speed, Overclock property.

CPU Core Count: It is the number of hardware chipsets in CPU.

Processor Bus Speed: It describes the speed of data processing. The speed is defined in terms of Megahertz.

Cache Memory Capacity: It is designed to speed up the transfer of data and instructions in terms of mega byte value.

CPU Speed: It is the speed which can be described in terms of frequency value that is denoted as Gigahertz at a microprocessor which executes instructions.

Maximum Clock Frequency: It provides to run a microprocessor faster than the speed for which it has been tested and approved. There are limit values that are denoted as Gigahertz.

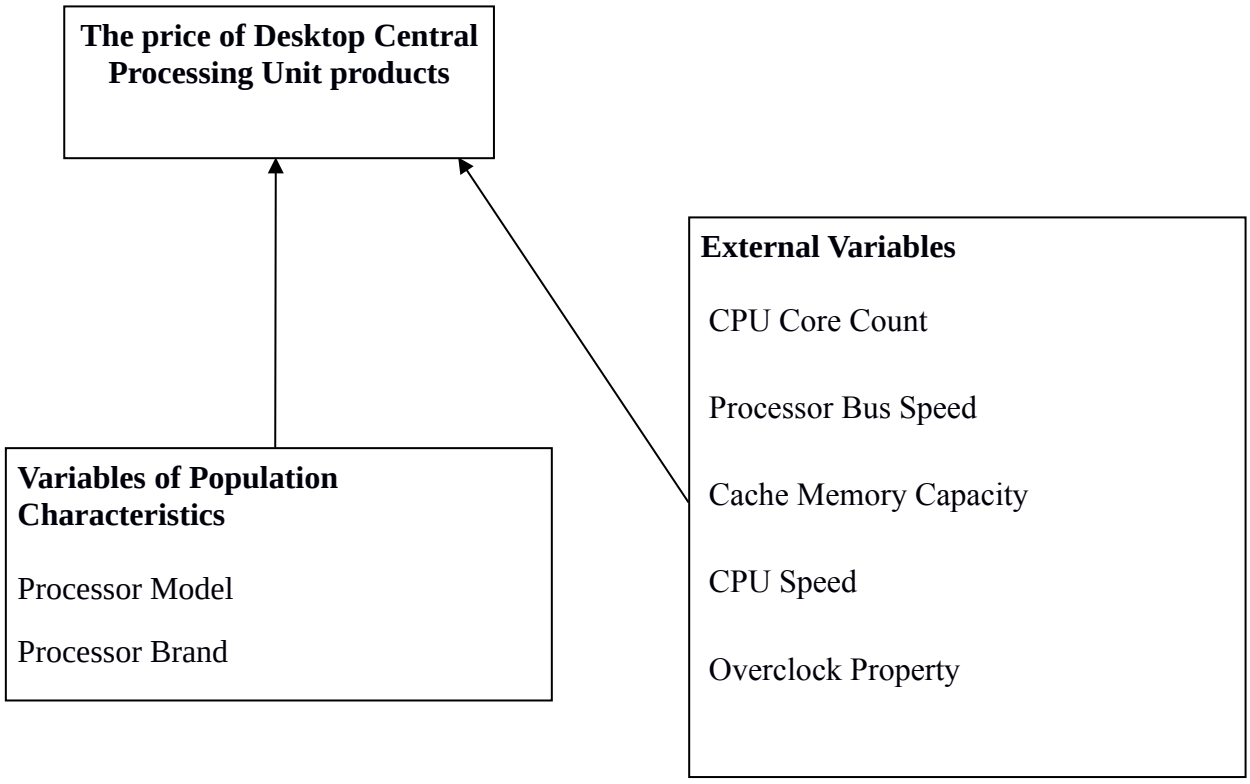
5. List of Variables for the characteristics of target population

Variables of population characteristics are Processor Model, Processor Brand.

Processor Model: It describes the overall architecture of CPU. It depends on the producers of CPU.

Processor Brand: It states the name of company which produces CPU products.

6. Core Concept Diagram



7. Methodology

7.1. Multiple Regression and Correlation Analysis

There are 7 steps for Multiple Regression and Correlation Analysis:

1. Linearity: Regression analysis is the method of analysis used to measure the relationship between two or more variables. If there is only one variable, it can use the same word to change them, it is a very variable regression.

2. MultiCollinearity: Multicollinearity is a very high correlation between at least two variables that predict a variant.

3. Normality: Normality tests are the examination of the normal distribution fitness of a given data set.

4. Homoscedastic: The variance of error terms is fixed, meaning that they do not change. The state of invalid is called the varying variance.

5. Adjusted r2: In a multiple regression analysis it is the adjusted coefficient of determination for the number of independent variables.

6. Autocorrelation: It is a correlation coefficient between the members of a time series observation.

7. Beta Coefficients: It corresponds to the standard regression coefficient. Beta is also known as weight. It explains the relative effect of each independent variable on the dependent variable.

7.2. Analysis of variance

Analysis of variance, also known as ANOVA, is perhaps the most powerful statistical tool. ANOVA is a general method of analyzing data from designed experiments, whose objective is to compare two or more group means. Whether the difference between the means of the two groups is meaningful can also be examined using the t test. If more than one group is to be compared, the F-Test is used as a variance analysis.

8. Findings

8.1. Multiple Regression and Correlation Analysis (MRCA)

Linearity

As a first step in MRCA, we tried to examine new external variables' relationship with core concept in terms of linearity. Only "corecount" has a linear relationship with core concept. The other variables have non-linear relationship. For that reason, we have to eliminate variables which non-linear relationship from model.

From Correlations table, we can see Pearson Correlation values which are related with core concept:

Pearson Correlation value of corecount = 0.883 >= 0.70 so it has linear relationship with core concept.

MultiCollinearity

In our case, we have only one variable so we don't have to look correlation between variables.

Homoscedasticity

We have one new variable in our model, so we have to create one residual values for this new variable. After creating residual variable, we have to create absolute residual values to use in Homoscedasticity analysis.

From Correlations table:

We should look the value of sig which is for between ABS_RES1 and corecount.

sig = 0.191 > alpha = 0.05 (It shows that model is homoscedastic.)

From this result, we can accept H0 hypothesis. We have a constant variance. It means that no relation between independent variable and the absolute value of error terms that's why the model is homoscedastic.

Normality of Error Terms

In this step, we try to understand whether the data distribution is normal or not.

From Tests of Normality table:

We can look for Kolmogorov-Smirnov section. Because, our data size is less than 1000.

sig = 0.200 >= 0.05

We can accept H0 hypothesis. It means that the distribution of data is normal.

Adjusted R Square and F-Test

From Model Summary Table:

Adjusted R Square = 0.772 >= 0.50

It shows that the model is powerful.

From ANOVA Table:

sig = 0.000 <= 0.05

Test is significant. We accepted H1 hypothesis. It shows that the model is valid.

Our model is significant due to the F-test result and 77.2% of the variations of dependent variable (price_cc) is explained on the variations in independent variable (corecount).

Auto Correlation Analysis

For auto correlation analysis Durbin-Watson table used for reference values. We have one component and 30 data records. So, k=1 and n=30. We used these information to determine “dl” and “du” values from Durbin-Watson Statistics table.

From Model Summary Table, we can use Durbin-Watson value as d=1.267

dl=1.134 and du=1.264 and d = 1.267 < 2

4 – du = 4 - 1. 264 = 2.736

du = 1. 264 < d=1. 267 < 4 – du = 2.736 (It shows that there is no autocorrelation.)

Test statistic value is in the range of autocorrelation region which implies that test result confirms the significance of the F-test that’s why the model is also considered as a valid model.

Beta Coefficient and T-Test

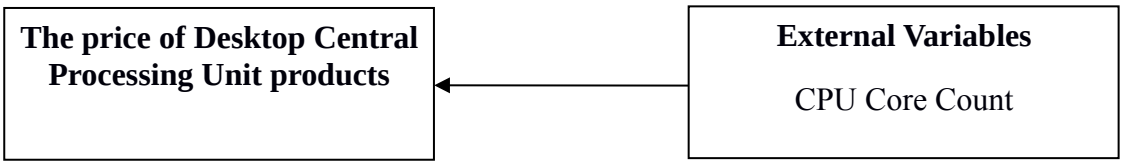
The beta values show the most effective variable is corecount.

From Coefficient Table:

sig of corecount = 0.000 <= 0.05

It shows that corecount independent variable has contribution to explanatory power.

8.2. Finalized Research Model based on MRCA



8.3. Analysis of variance (ANOVA)

Firstly, Levene's Test is applied in order to make sure, our data is satisfy assumption of ANOVA. sig=0.405 > alpha=0.05, so the variances of groups are homogeneous.

Levene's Test of Equality of Error Variances[a]

Dependent Variable: price_cc

F	df1	df2	Sig.
.715	1	28	.405

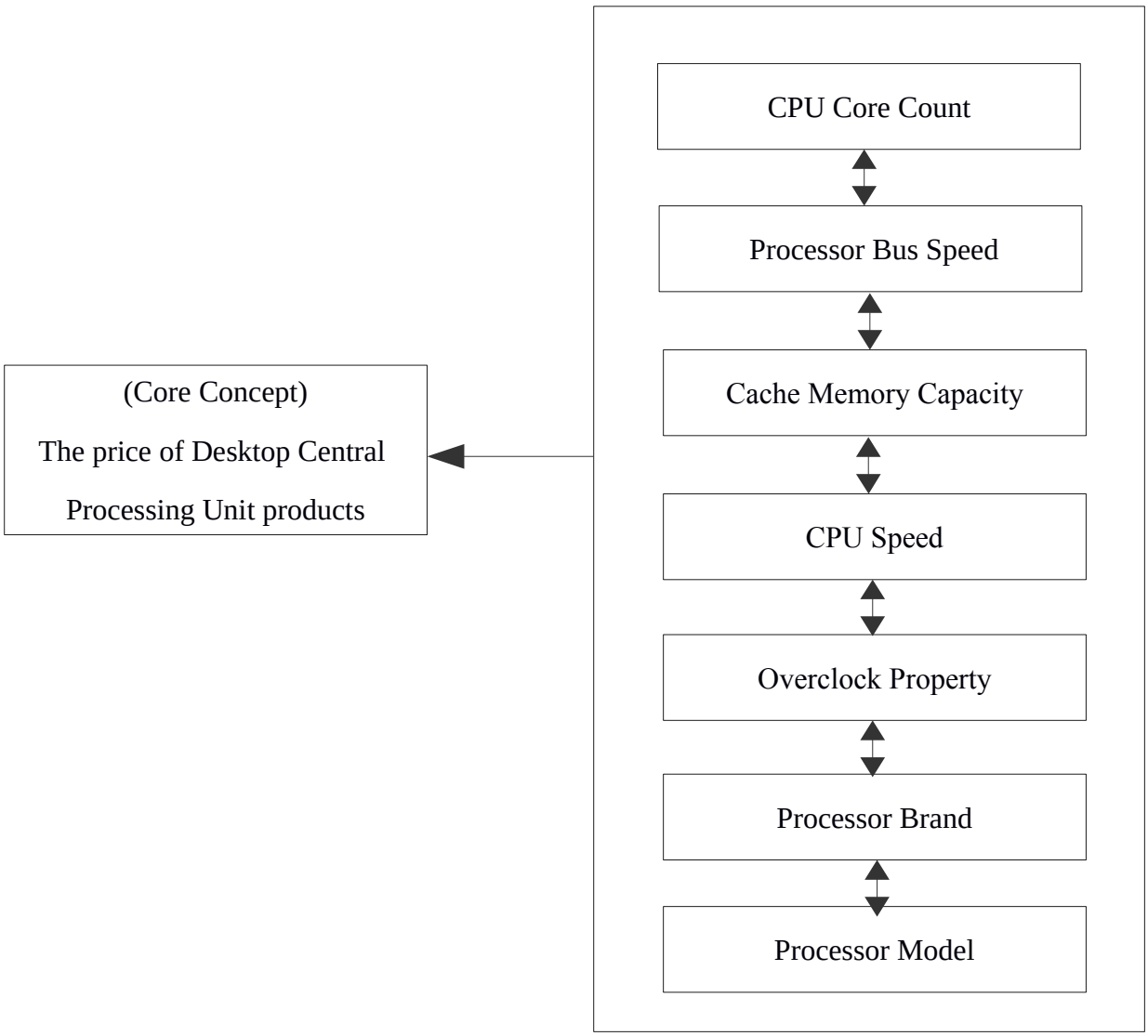
Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

We look up sources which have sig < α = 0.05. Our 7-way interaction have sig=0.000 with 0.995 observed power. That mean there is a significant 7-way interaction effect on our core concept.

sig=0.000 < 0.05

brand * corecount * bus_speed * cache_memory_capacity * cpu_speed * max_clock_freq * processor_model	2565183.93	2	1282591.97	14.058	.000	28.116	.995

8.4. Finalized Research Model based on ANOVA



9. Conclusion

Our ANOVA analysis show that price of Desktop version CPU is effected by core count, processor bus speed, cache memory capacity, speed, overclock property, brand and model. We examined Intel and AMD cpu products in this study. Even though all of named variables effects performance, the most effective factor is the core count of CPU.

10. Limitations

The variation of desktop CPU is limited in a small range. There are two big CPU producer company Intel and AMD. But, The volume of desktop CPU products in the market is small. So we can not increase the data size easily.

11. Implications

In further projects, Overclock Property can be replaced with another more powerful factor by researcher. Also, beside Intel and AMD companies, researchers can add other small desktop CPU producer companies to their research project.

12. References

[1] *Factors Affecting CPU Performance.* URL: `https://www.maketecheasier.com/factors-affecting-cpu-performance/`

[2] *The factors affecting the performance of the CPU.* URL: `http://theteacher.info/index.php/computing-principles-01/processors-and-components/notes/2066-the-factors-affecting-the-performance-of-the-cpu`

13. Appendix

13.1. Labeled & Coded-Questionnaire

1	corecount	bus_speed	cache_memory_capacity	cpu_speed	max_clock_freq	brand	processor_model	price_cc
2	4.00	2560.00	8.00	4.00	4.00	1.00	1.00	881.04
3	2.00	2560.00	3.00	3.60	3.60	1.00	2.00	401.65
4	16.00	2667.00	32.00	3.60	4.00	2.00	3.00	5135.05
5	4.00	2133.00	8.00	3.10	3.40	2.00	4.00	494.98
6	4.00	2667.00	8.00	3.50	3.70	2.00	4.00	596.86
7	4.00	2560.00	6.00	4.00	4.20	1.00	5.00	1075.38
8	4.00	2560.00	8.00	4.30	4.50	1.00	6.00	1576.36
9	10.00	2560.00	13.75	3.30	4.30	1.00	7.00	5005.49
10	6.00	2560.00	8.25	3.50	4.00	1.00	6.00	1675.49
11	8.00	2560.00	11.00	3.60	4.30	1.00	6.00	2647.56
12	4.00	3200.00	8.00	3.20	3.40	2.00	8.00	747.15
13	6.00	3200.00	16.00	3.40	3.60	2.00	8.00	965.69
14	6.00	3200.00	16.00	3.60	4.00	2.00	8.00	1120.45
15	4.00	3200.00	16.00	3.60	3.70	2.00	8.00	849.21
16	8.00	3000.00	20.00	3.40	3.80	2.00	9.00	1626.41
17	4.00	2400.00	6.00	3.00	3.50	1.00	5.00	794.01
18	4.00	2400.00	6.00	3.40	3.80	1.00	5.00	885.56
19	2.00	2400.00	3.00	3.90	3.90	1.00	1.00	505.78
20	4.00	2400.00	8.00	4.20	4.50	1.00	6.00	1524.44
21	6.00	2400.00	15.00	3.60	3.80	1.00	6.00	2245.61
22	4.00	2133.00	4.00	3.60	4.00	2.00	10.00	358.46
23	2.00	2133.00	3.00	3.30	3.30	1.00	2.00	267.11
24	2.00	2133.00	6.00	3.70	3.70	1.00	1.00	477.66
25	4.00	2133.00	6.00	2.70	3.30	1.00	5.00	846.88
26	4.00	0.00	4.00	3.10	3.80	2.00	11.00	237.01
27	4.00	2133.00	4.00	3.30	3.80	2.00	12.00	375.65
28	4.00	2133.00	4.00	3.10	3.80	2.00	12.00	298.00
29	8.00	0.00	16.00	4.70	5.00	2.00	13.00	729.08
30	2.00	1866.00	1.00	3.90	4.10	2.00	14.00	159.77
31	12.00	2560.00	16.50	2.90	4.30	1.00	7.00	5610.12

13.2. Outputs of Analysis

13.2.1. Outputs of MRCA

Regression

Descriptive Statistics

	Mean	Std. Deviation	N
price_cc	1337.1303	1455.01148	30
corecount	5.2000	3.13380	30

Correlations

		price_cc	corecount
Pearson Correlation	price_cc	1.000	.883
	corecount	.883	1.000
Sig. (1-tailed)	price_cc	.	.000
	corecount	.000	.
N	price_cc	30	30
	corecount	30	30

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	corecount[b]	.	Enter

a. Dependent Variable: price_cc

b. All requested variables entered.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.883[a]	.780	.772	694.18814	1.267

a. Predictors: (Constant), corecount

b. Dependent Variable: price_cc

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	47901573.1	1	47901573.1	99.402	.000[b]
	Residual	13493121.0	28	481897.178		
	Total	61394694.1	29			

a. Dependent Variable: price_cc

b. Predictors: (Constant), corecount

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	-795.464	248.629		-3.199	.003
	corecount	410.114	41.135	.883	9.970	.000

a. Dependent Variable: price_cc

Residuals Statistics[a]

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	24.7646	5766.3647	1337.1303	1285.21531	30
Residual	-1756.3704	1699.81116	.00000	682.11438	30
Std. Predicted Value	-1.021	3.446	.000	1.000	30
Std. Residual	-2.530	2.449	.000	.983	30

a. Dependent Variable: price_cc

Correlations

			price_cc	ABS_RES1
Spearman's rho	price_cc	Correlation Coefficient	1.000	.245
		Sig. (2-tailed)	.	.191
		N	30	30
	ABS_RES1	Correlation Coefficient	.245	1.000
		Sig. (2-tailed)	.191	.
		N	30	30

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	.095	30	.200*	.960	30	.305

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

13.2.2. Outputs of ANOVA

Univariate Analysis of Variance

Between-Subjects Factors

		N
brand	1.00	16
	2.00	14

Levene's Test of Equality of Error Variances[a]

Dependent Variable: price_cc

F	df1	df2	Sig.
.715	1	28	.405

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + brand * corecount * bus_speed * cache_memory_capacity * cpu_speed * max_clock_freq * processor_model + brand + corecount + bus_speed + cache_memory_capacity + cpu_speed + max_clock_freq + processor_model

Tests of Between-Subjects Effects

Dependent Variable: price_cc

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Noncent. Parameter	Observed Power[b]
Corrected Model	59569953[a]	9	6618883.70	72.546	.000	652.914	1.000
Intercept	2527.324	1	2527.324	.028	.869	.028	.053
brand * corecount * bus_speed * cache_memory_capacity * cpu_speed * max_clock_freq * processor_model	2565183.93	2	1282591.97	14.058	.000	28.116	.995
brand	4037.121	1	4037.121	.044	.836	.044	.055
corecount	2472799.15	1	2472799.15	27.103	.000	27.103	.999
bus_speed	22175.837	1	22175.837	.243	.627	.243	.076
cache_memory_capacity	237107.653	1	237107.653	2.599	.123	2.599	.336
cpu_speed	150147.015	1	150147.015	1.646	.214	1.646	.231
max_clock_freq	98516.722	1	98516.722	1.080	.311	1.080	.168
processor_model	159117.956	1	159117.956	1.744	.202	1.744	.242
Error	1824740.72	20	91237.036				
Total	115032220	30					
Corrected Total	61394694.1	29					

a. R Squared = .970 (Adjusted R Squared = .957)

b. Computed using alpha = .05