**Applied Machine Learning Project Report**

**Hakkıcan BÜLÜÇ - 180709061 Adem VAROL - 200709078 / Hüseyin ÇOBAN - 180709018**

**CHAPTER 1: INTRODUCTION**

**1-A. Overview:**

In this research, we examined the students of Ortaca Anatolian High School (OAL) and Yunus Emre Anatolian High School (YEAL). In this research, we investigate the question that most affects students' success. Our research questions:

1. Are students living in the village more successful than those living in the city center?
2. Are girls or boys most negatively affected by the separation of parents?
3. Which attribute is more important to success?
4. Can a new student get a passing grade?
5. At which grade level is the educational status of the parents more effective?

**1-B. Statement of the Problem:**

In our research, we explore the question that most affects the success of high school students. The reason we are investigating this is that if we can find the factor that affects students the most, we can go after it and try to solve the problem. In our research, apart from the question that most affected the success of high school students, other questions we investigated:

1. Are students living in the village more successful than those living in the city center?
2. Are girls or boys most negatively affected by the separation of parents?
3. Can a new student get a passing grade?
4. At which grade level is the educational status of the parents more effective?

**1-C. Purpose**

Our aim is to determine the factor that most affects the success of Ortaca Anatolian High School (OAL) and Yunus Emre Anatolian High School (YEAL) students. We will try to achieve this goal by starting with the exploratory questions we have set to realize our goal. We will test our hypotheses. You can find our hypotheses in chapter 2.

**1-D.1. Research Questions or Hypotheses**

Our research questions:

1. Are students living in the village more successful than those living in the city center?
2. Are girls or boys most negatively affected by the separation of parents?
3. Which attribute is more important to success?
4. Can a new student get a passing grade?
5. At which grade level is the educational status of the parents more effective?

Our hypotheses:

1. We think that students living in the city center are more successful. Because we think that students living in the city center will be more successful because they can meet their needs more easily.
2. We think that the separation of the parents will affect the girls most negatively. Because we think that girl students are more emotional.
3. The most important question for success is "Is the father graduated from primary school at most?”. Because students whose fathers are educated are also educated.
4. Since our average is low, we think that if a new student is added to our data, he/she will fail.
5. We think that the education level of the parents will affect the 12th grades the most. Because students whose parents are educated work harder because they want to be successful like their parents.

**1-D.2. Significance of the Study:**

BÖLÜM ÇIKARMAYALAIM. AÇIKLAMA YUKARIDA YAPILMIŞTIR GİBİ BİR ŞEYLER YAZALIM BENCE.

**1-E. Conceptual Framework:**

BÖLÜM ÇIKARMAYALAIM. AÇIKLAMA YUKARIDA YAPILMIŞTIR GİBİ BİR ŞEYLER YAZALIM BENCE.

|  |  |
| --- | --- |
| *Quantitative Research:*  *Research Questions and/or Hypotheses* | *Qualitative Research:*  *Foreshadowed Problems, Conjectures,bor Exploratory Questions* |
| Our hypotheses:   1. We think that students living in the city center are more successful. Because students living in the city center are more successful because they can meet their needs more easily. 2. We think that the separation of the parents will affect the girls most negatively. Because girls are more emotional. 3. The most important question for success is "Is the father graduated from primary school at most?". Because students whose fathers are educated are also educated. 4. Since our average is low, we think that if a new student is added to our data, they will fail. 5. We think that the education level of the parents will affect the 12th grades the most. Because students whose parents are educated work harder because they want to be successful like their parents. | Our exploration questions:   1. Are students living in the village more successful than those living in the city center? 2. Are girls or boys most negatively affected by the separation of parents? 3. Which attribute is more important to success? 4. Can a new student get a passing grade? 5. At which grade level is the educational status of the parents more effective? |

We used the R statistical computing and graphical programming language to do our research. We knew the simple basics about this programming language. To answer our questions, we had to learn the advanced level of the programming language.

**1-F**. **Summary of Methodology**:

Our problem is to determine the factor that most affects the success of high school students. We analyzed the data with R statistical computing and graphical programming language. We used some packages of the R programming language to answer some of our questions. Packages we use: readxl, data.table, FSelectorRcpp, e1071, caTools, caret, dplyr.

**1-G**. **Limitations**:

Our biggest limitation is that we cannot conduct this survey to all high school students in our country. For this reason, we applied this survey only to nearby high schools. The reason why we could not conduct this survey to all high school students in our country is due to the lack of resources.

**1-H. Definition of Terms:**

* **Advanced:** At a higher, more difficult level.
* **Affect:** To have an influence on someone or something, or to cause a change in someone or something.
* **Attribute:** A quality or feature of a person or thing, esp. one that is an important part of its nature.
* **Conduct:** To organize and perform a particular activity.
* **Demographic:** A group of people, for example customers, who are similar in age, social class, etc.
* **Excel:** A brand of software used for creating spreadsheets (= documents with rows of data for showing amounts and making calculations).
* **Exploratory:** Done in order to discover more about something.
* **Factor:** A fact or situation that influences the result of something.
* **Grade:** A school class or group of classes in which all the children are of a similar age or ability.
* **Hypothesis:** An idea or explanation for something that is based on known facts but has not yet been proved.
* **Influence:** The power to have an effect on people or things, or a person or thing that is able to do this.
* **Needs:** The things you must have for a satisfactory life.
* **Programming:** The instructions that tell a computer what to do.
* **R Studio:** Is an integrated development environment (IDE) for R.
* **Tentatively:** In a way that shows you are not certain or confident.

**CHAPTER 2 REVIEW OF THE LITERATURE**

BU SORULAR HER YENİ GELEN ÖĞRENCİYE SORULUYOR AMA BU SORULAR ÜZERİNDE BÖYLE BİR ÇALIŞMA YAPILMADI TABİKİ DE…

In this research, we examined the students of Ortaca Anatolian High School (OAL) and Yunus Emre Anatolian High School (YEAL). The students were asked about their demographic characteristics, the education level of their parents, the number of siblings, the divorce and separation status of the parents, who the student lived with, the living conditions of his parents, and whether there was any illness in his family or himself. No such research has been done for OAL and YEAL schools before. Therefore, this research is a first.

Our research questions:

1. Are students living in the village more successful than those living in the city center?
2. Are girls or boys most negatively affected by the separation of parents?
3. Which attribute is more important to success?
4. Can a new student get a passing grade?
5. At which grade level is the educational status of the parents more effective?

Our hypotheses:

1. We think that students living in the city center are more successful. Because we think that students living in the city center will be more successful because they can meet their needs more easily.
2. We think that the separation of the parents will affect the girls most negatively. Because we think that girl students are more emotional.
3. The most important question for success is "Is the father graduated from primary school at most?”. Because students whose fathers are educated are also educated.
4. Since our average is low, we think that if a new student is added to our data, he/she will fail.
5. We think that the education level of the parents will affect the 12th grades the most. Because students whose parents are educated work harder because they want to be successful like their parents.

**2-A. Introduction: Topic(s), Purposes, and Methods of the Literature Review:**

In this research, we investigate the question that most affects students' success. We analyzed the data with R statistical computing and graphical programming language.

**2-B. Description and Critique of Scholarly Literature:**

The strength of our research is that we learned the demographic characteristics of our students. The weakness is that the students are asked a limited number of questions, so the results are only as strong as the data we have.

**2-C. Inferences for Forthcoming Study:**

BÖLÜM ÇIKARMAYALAIM. AÇIKLAMA YUKARIDA YAPILMIŞTIR GİBİ BİR ŞEYLER YAZALIM BENCE.

**2-D. Theoretical/Conceptual Framework for Forthcoming Study (May appear in chapter 3).**

BÖLÜM ÇIKARMAYALAIM. AÇIKLAMA YUKARIDA YAPILMIŞTIR GİBİ BİR ŞEYLER YAZALIM BENCE.

**CHAPTER 3: RESEARCH DESIGN AND METHODOLOGY**

**3-A. Methodology:**

We analyzed the data with the R programming language. We used the packages of the R programming language to find answers to our questions. Packages we use:

1. readxl: With this package, we transferred our data in excel to R Studio.
2. data.table: With this package, we filtered our data frame with the like keyword.
3. FSelectorRcpp: We calculated information gain with this package.
4. e1071, caTools, caret: With these packages, we estimated the success of a new student.
5. dplyr. With this package, we filtered our data frame.

**3-B. Conjectures, or Exploratory Questions:**

Our exploration questions:

1. Are students living in the village more successful than those living in the city center?
2. Are girls or boys most negatively affected by the separation of parents?
3. Which attribute is more important to success?
4. Can a new student get a passing grade?
5. At which grade level is the educational status of the parents more effective?

**3-C. Research Procedures:**

Bu araştırmamız kişilere ulaşmanın daha kolay olduğunu düşündüğümüz için elektronik ortamdaki formlar üzerinden yapıldı ve bilgisayar ortamındaki tablolarda tutuldu.. Herhangi bir ses veya görüntü kaydı alınmadı. Bu veriler üzerinde daha önce herhangi bir analiz yapılmadı.

**3-D. Human Participants and Ethics Precautions:**

Araştırmaya katılan kişiler kendi rızaları ile bu bilgileri paylaşmışlardır. Araştırmamızdaki sorulara gerçekçi yanıt vermeleri için bilgilerinin hiç kimse ile paylaşılmayacağı kendilerine ifade edilmiştir. Ancak bilgileri aldığımız kurum/kurumlar araştırmaya katılan kişilerin isimlerinin paylaşılmaması şartıyla bu verileri bizimle paylaşmışlardır. Zaten yapacağımız veri analizinde katılımcıların isimlerinin bir önemi yoktur.

Analizimizi tamamladığımızda analiz sonucumuz kurum/kurumlar ile paylaşılarak yapılan anketlerin eksik veya fazla yönleri detaylıca tartışılacaktır. Bu analizimiz doğrultusunda yapılan anket soruları tekrar gözden geçirilerek eklemeler veya çıkarmalar yapılacaktır.

**CHAPTER 4: RESULTS or FINDINGS**

**Our exploration questions and answers:**

1. Are students living in the village more successful than those living in the city center?

Students living in the city center are more successful.

1. Are girls or boys most negatively affected by the separation of parents?

The students most affected by the separation of mothers and fathers are boy students.

1. Which attribute is more important to success?

Baba.en.fazla.ilkokul.mezunu.mu.

1. Can a new student get a passing grade?

Confusion Matrix’s 'Positive' Class is 0.

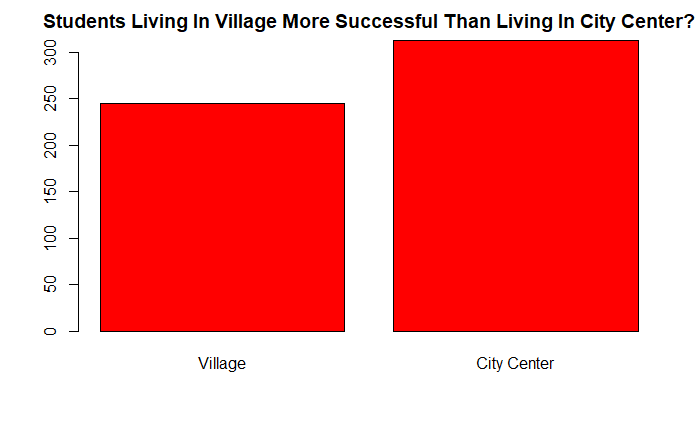
1. At which grade level is the educational status of the parents more effective?

The education level of the parents affected the 12th grades the most.

**Bar graphs of some of our questions:**

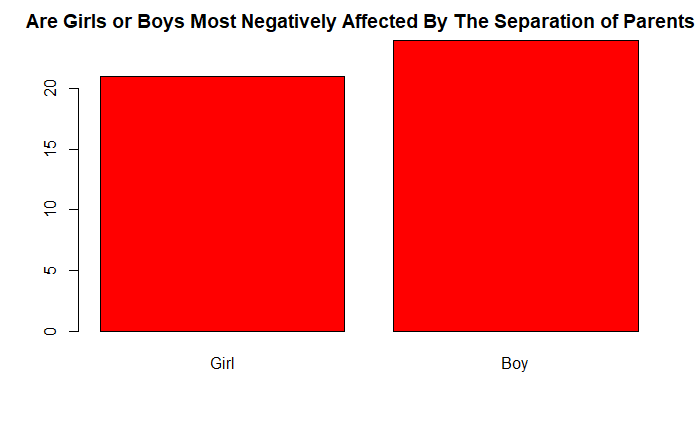
1. Are students living in the village more successful than those living in the city center?

As you can see in the graphic below, the number of successful students living in the city center is higher than those living in the village.



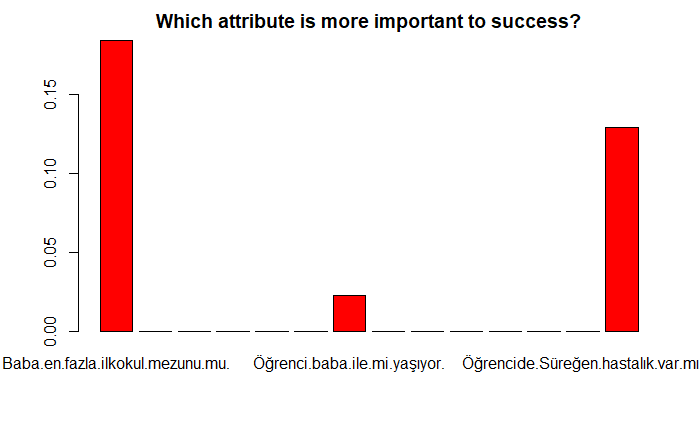
1. Are girls or boys most negatively affected by the separation of parents?

As you can see in the graphic below, the number of boy students who are negatively affected by the separation of their parents is higher than that of girl students.



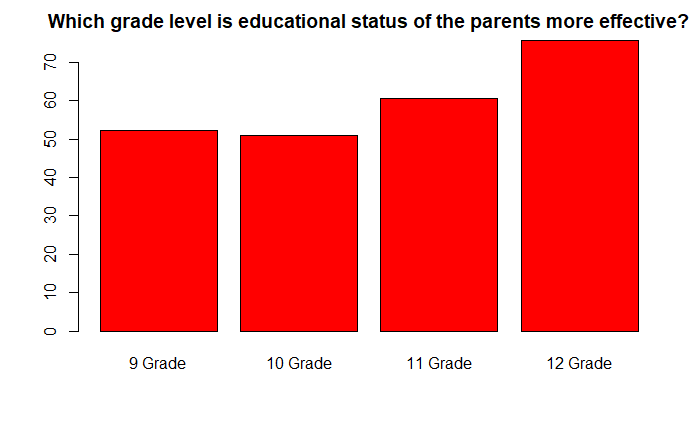
1. Which attribute is more important to success?

In the graphic below, you can see the importance of our survey questions. As you can see in the graphic, the most important question is Baba.en.fazla.ilkokul.mezunu.mu.



5. At which grade level is the educational status of the parents more effective?

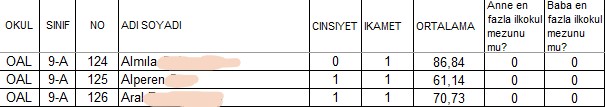
As you can see in the graphic below, the education level of the parents affects the 12th grades the most.



**4-A. Text:**

The most important question of our research is which question is most important to student success. The answer of our research is Baba.en.fazla.ilkokul.mezunu.mu. . As a result of our research, oal school students are more successful.

**4-B. Raw Data:**

****

**CHAPTER 5: CONCLUSIONS, INTERPRETATIONS AND RECOMMENDATIONS**

Our exploration questions, answers, and comments:

1. Are students living in the village more successful than those living in the city center?

Students living in the city center are more successful. Because, in the data we have, such a result was obtained because the students living in the city center are the most.

1. Are girls or boys most negatively affected by the separation of parents?

The students most affected by the separation of parents are boy students. Because there are mostly boy students in the data we have. As this answer reveals, boys can be just as emotional as girls.

1. Which attribute is more important to success?

Baba.en.fazla.ilkokul.mezunu.mu. The oal and yeal schools where our data were collected are in Muğla / Ortaca. Generally, educated fathers living in Ortaca make their own children educated.

1. Can a new student get a passing grade?

Confusion Matrix’s 'Positive' Class is 0. Since the average of our students is low, if a new student joins these students, our existing unsuccessful students will be unsuccessful as it will affect our new student.

1. At which grade level is the educational status of the parents more effective?

The education level of the parents affected the 12th grades the most. Although there were at least 12th grade students in our data, 12th grade students were the most affected. The reason why 12th grade students are the most affected by this situation is that the parents of the students whose parents are educated support their children.

**5-A. Summary:**

In our research, we investigated the question that affects the success of high school students the most. As a result of our research, the fact that our students' fathers were primary school graduates at most was the question that most affected the success of our students.

Our hypotheses:

1. We think that students living in the city center are more successful. This hypothesis was supported.
2. We think that the separation of the parents will affect the girls most negatively. This hypothesis was not supported.
3. The most important question for success is "Is the father graduated from primary school at most?". This hypothesis was supported.
4. Since our average is low, we think that if a new student is added to our data, they will fail. This hypothesis was supported.
5. We think that the education level of the parents will affect the 12th grades the most. This hypothesis was supported.

**5-B. Conclusions:**

Although our research results have the meanings mentioned above locally, the results do not cover the whole country, as our survey study was not conducted throughout the country. If our results are worked on, we think that the success of high school students will increase first locally and then across the country. The results of our research show that, as far as the accuracy of our data is, the factor that most affected the success of high school students was that the fathers of the students were primary school graduates the most.

As far as we learned from our research, it was boy students who were most affected by the separation of parents. Such a result was obtained because of our data. This result is tentatively recommended.

As a result of our research, high school students living in the city center may be more successful.

**5-C. Recommendations:**

As a result of our research, the most influential factor in the success of high school students is the educational status of the students' fathers. Therefore, the education of future generations depends on the education of men. The education of boy students should be given importance.

As we learned from the research, students living in the city center are more successful. Therefore, investments made in the field of education in villages can be increased in order to increase the success of students living in villages.

Boy students are most negatively affected by the separation of their parents. Therefore, more psychological support can be provided to boys after the parent’s separation.

According to the data we have, our students are unsuccessful. Therefore, a new student should not be joined to these students.

As a result of our research, we have seen that 12th grade students are the most affected by the educational status of parents. Therefore, the education of present children should be given importance for the success of their future children.