



**NİŞANTAŞI**  
**ÜNİVERSİTESİ**

Veri Madenciliği Dersi  
Final Ödevi

Kadir BERKTAŞ  
**20239001068**

## **İçindekiler:**

|  |           |
|--|-----------|
| <b>Veri Seti Hakkında:</b> .....                       | <b>3</b>  |
| • Veri Seti Amacı ve İçeriği                           |           |
| • Veri Setinin Kaynağı                                 |           |
| • Veri Setindeki Değişkenler                           |           |
| • Verinin görüntülenmesi                               |           |
| <b>Veri Ön İşleme:</b> .....                           | <b>5</b>  |
| • Verinin ön işlenmesi                                 |           |
| <b>Veri Seti Keşfi ve Görselleştirilmesi :</b> .....   | <b>6</b>  |
| • Demografik Bilgiler ve Finansal Durum                |           |
| • Kampanya Sonuçları ve Müşteri Özellikleri            |           |
| • Bankacılık İlişkisi ve Finansal Durum                |           |
| • Kampanya Performansı                                 |           |
| • Zaman Bazlı Analiz                                   |           |
| • Müşteri İletişim Türü                                |           |
| • Önceki Kampanya Sonuçları ve Mevcut Durum            |           |
| <b>Model Seçimi ve Eğitim:</b> .....                   | <b>14</b> |
| • K-Means Algoritmanın Uygulanması                     |           |
| <b>Sonuçların Yorumlanması ve Analiz :</b> .....       | <b>15</b> |
| • Müşteri segmentlerinin görselleştirerek yorumlanması |           |
| • Müşteri segmentlerinin özet istatistikleri           |           |
| <b>Kaynak:</b> .....                                   | <b>19</b> |
| <b>Python Kodları:</b> .....                           | <b>20</b> |

## Veri Setinin Amacı ve İçeriği

Bu veri seti, bir bankanın doğrudan pazarlama kampanyaları sırasında topladığı müşteri bilgilerini içerir. Özellikle, bankanın vadeli mevduat ürünlerini pazarlamak amacıyla gerçekleştirdiği telefon görüşmeleri sonucunda elde edilen veriler bulunmaktadır. Amaç, hangi müşteri gruplarının bu kampanyalara daha olumlu tepki verdiğini anlamak ve gelecekteki pazarlama stratejilerini optimize etmektir.

## Veri Setinin Kaynağı

Veriler, bir Avrupa ülkesindeki banka tarafından toplanmıştır. Banka, kampanya sırasında müşterilerine telefonla ulaşmış ve bu görüşmelerden çeşitli müşteri özelliklerini ve kampanya sonuçlarını kaydetmiştir. Veri setine “Kaggle” üzerinden ulaşılmıştır. Kaynak kısmında linki bulabilirsiniz.

## Veri Setindeki Değişkenler

- **age:** Müşterinin yaşı (tam sayı).
- **job:** Müşterinin mesleği (metinsel veri).
- **marital:** Müşterinin medeni durumu (metinsel veri).
- **education:** Müşterinin eğitim durumu (metinsel veri).
- **default:** Müşterinin kredi kartı borcunu ödeyip ödemediği durumu (metinsel veri).
- **balance:** Müşterinin banka hesabındaki mevcut bakiyesi (tam sayı).
- **housing:** Müşterinin konut kredisine sahip olup olmadığı (metinsel veri).
- **loan:** Müşterinin kişisel krediye sahip olup olmadığı (metinsel veri).
- **contact:** Müşteriyle iletişim kurulan iletişim türü (metinsel veri).
- **day:** İletişim kurulan gün (tam sayı).
- **month:** İletişim kurulan ay (metinsel veri).
- **duration:** İletişimin süresi (tam sayı, saniye cinsinden).
- **campaign:** Mevcut kampanya sırasında müşteriyle yapılan görüşme sayısı (tam sayı).
- **pdays:** Müşteriyle en son yapılan kampanyadan bu yana geçen gün sayısı (tam sayı).
- **previous:** Müşteriyle yapılan önceki kampanya sayısı (tam sayı).
- **poutcome:** Önceki kampanyanın sonucu (metinsel veri).
- **deposit:** Müşterinin vadeli mevduat teklifini kabul edip etmediği (metinsel veri, hedef değişken).

Verideki ilk beş satır aşağıdaki gibi görüntülenmektedir:

|   | age | job        | marital | education | default | balance | housing | loan | contact | \ |
|---|-----|------------|---------|-----------|---------|---------|---------|------|---------|---|
| 0 | 59  | admin.     | married | secondary | no      | 2343    | yes     | no   | unknown |   |
| 1 | 56  | admin.     | married | secondary | no      | 45      | no      | no   | unknown |   |
| 2 | 41  | technician | married | secondary | no      | 1270    | yes     | no   | unknown |   |
| 3 | 55  | services   | married | secondary | no      | 2476    | yes     | no   | unknown |   |
| 4 | 54  | admin.     | married | tertiary  | no      | 184     | no      | no   | unknown |   |

|   | day | month | duration | campaign | pdays | previous | poutcome | deposit |
|---|-----|-------|----------|----------|-------|----------|----------|---------|
| 0 | 5   | may   | 1042     | 1        | -1    | 0        | unknown  | yes     |
| 1 | 5   | may   | 1467     | 1        | -1    | 0        | unknown  | yes     |
| 2 | 5   | may   | 1389     | 1        | -1    | 0        | unknown  | yes     |
| 3 | 5   | may   | 579      | 1        | -1    | 0        | unknown  | yes     |
| 4 | 5   | may   | 673      | 2        | -1    | 0        | unknown  | yes     |

## Veri Ön İşleme

Veri setinin özet bilgilerinin görüntülenmesi:

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 11162 entries, 0 to 11161
Data columns (total 17 columns):
#   Column      Non-Null Count  Dtype
---  -
0   age         11162 non-null  int64
1   job         11162 non-null  object
2   marital     11162 non-null  object
3   education   11162 non-null  object
4   default     11162 non-null  object
5   balance     11162 non-null  int64
6   housing     11162 non-null  object
7   loan        11162 non-null  object
8   contact     11162 non-null  object
9   day         11162 non-null  int64
10  month       11162 non-null  object
11  duration    11162 non-null  int64
12  campaign    11162 non-null  int64
13  pdays      11162 non-null  int64
14  previous    11162 non-null  int64
15  poutcome    11162 non-null  object
16  deposit     11162 non-null  object
dtypes: int64(7), object(10)
memory usage: 1.4+ MB
None
```

Eksik değerlerin olup olmadığının kontrol edilmesi:

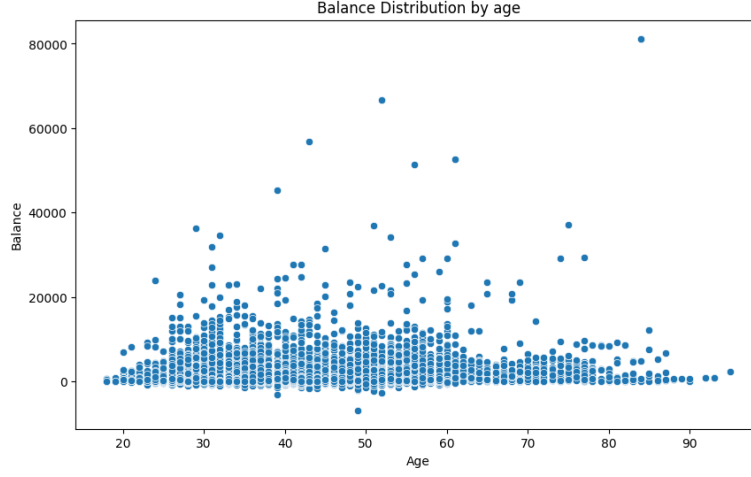
```
age         0
job         0
marital     0
education   0
default     0
balance     0
housing     0
loan        0
contact     0
day         0
month       0
duration    0
campaign    0
pdays      0
previous    0
poutcome    0
deposit     0
dtype: int64
```

**Veri setinde eksik değer olmadığından dolayı, veri üzerinde gerekli keşif ve görselleştirmeler yapılabilir.**

# Veri Keşfi ve Görselleştirme

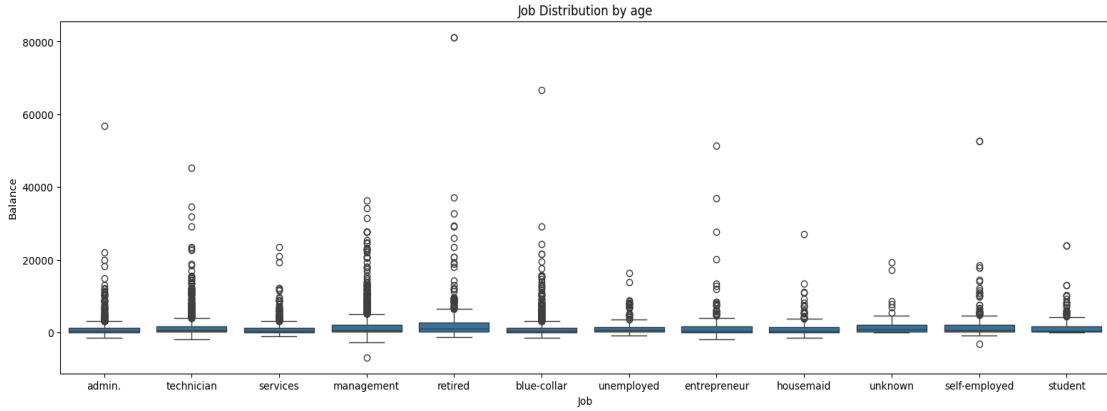
## Demografik Bilgiler ve Finansal Durum

- **Age ve balance:** Müşterilerin yaş grupları ile banka bakiyeleri arasında bir ilişki olup olmadığını aşağıda görüntülenmektedir. Burada sorulan soru; Yaş arttıkça bakiye değişiklik gösteriyor mu?



Grafikten çıkardığım bilgilere göre yaş arttıkça müşterilerin banka bakiyeleri arasında net bir artışın olduğu söylenemez. Hatta düşük oranda azalışın olduğu söylenebilir.

- **Job ve balance:** Farklı meslek gruplarının bankadaki bakiyeleri aşağıda görüntülenmektedir.

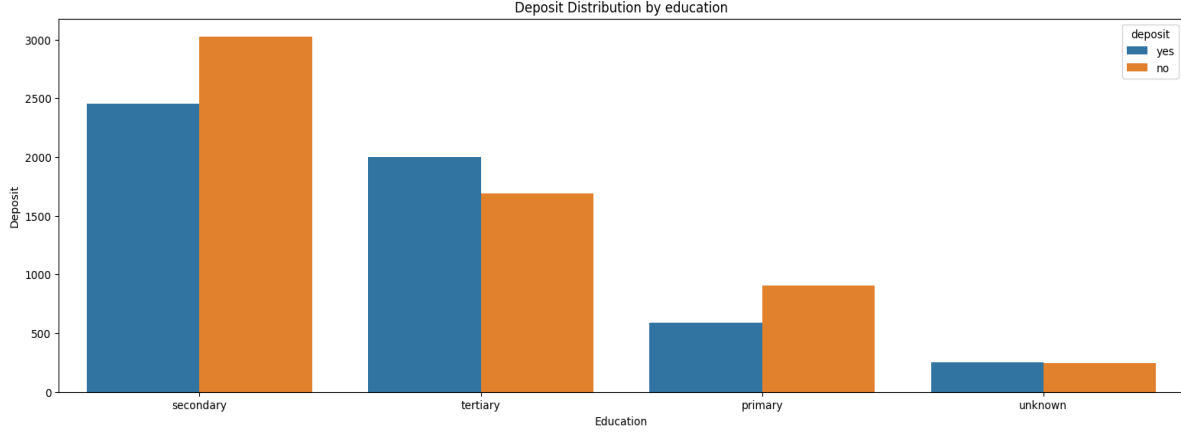


**Teknisyenlerde, yöneticilerde ve emeklilerdeki** banka bakiyeleri diğer meslek gruplarına göre daha yüksektir.

**İşsizler, kendi işinin sahibi olanlar ve öğrencilerdeki** banka bakiyeleri diğer meslek gruplarına göre daha düşüktür.

## Kampanya Sonuçları ve Müşteri Özellikleri:

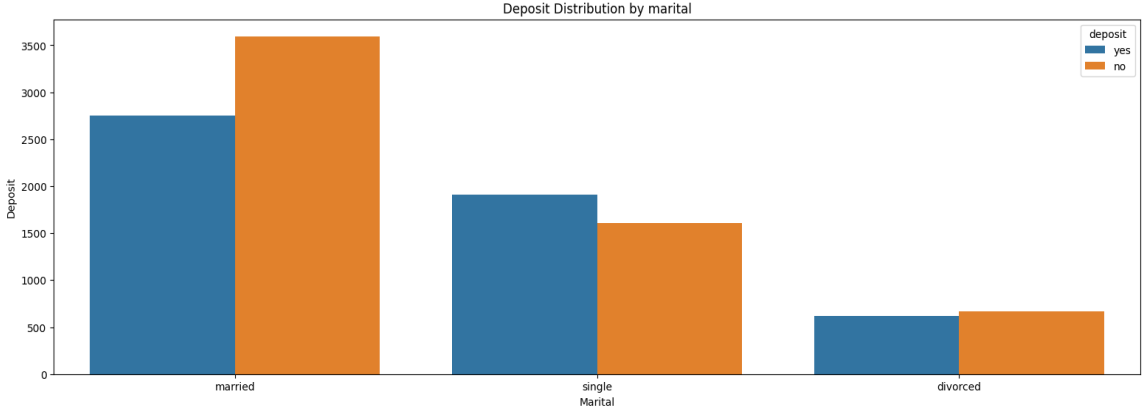
- **Education ve deposit:** Eğitim düzeyinin vadeli mevduat teklifine olumlu yanıt verme olasılığı üzerindeki etkisinin grafiği aşağıda görüntülenmektedir.



**Ortaöğretim mezunu ve yükseköğrenim mezunu** kişiler vadeli mevduat tekliflerine olumlu ve olumsuz cevap verme oranları en **yüksektir**.

**İlköğretim mezunu ve eğitim durumu** bilinmeyen kişiler ise vadeli mevduat tekliflerine olum ve olumsuz cevap verme oranları en **düşüktür**.

- **Marital ve deposit:** Medeni durumun, müşterilerin vadeli mevduat kabul etme oranları üzerindeki etkisinin grafiği aşağıda görüntülenmektedir.

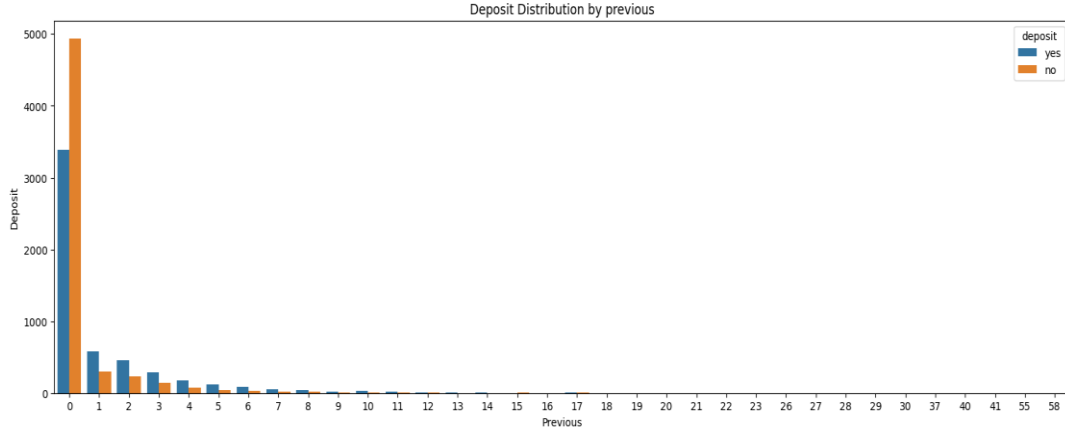


**Evli** insanların vadeli mevduat tekliflerine olumlu ve olumsuz cevap verme oranı en yüksektir

**Bekar** insanların vadeli mevduat tekliflerine olumlu ve olumsuz cevap verme oranı evli insanlara göre daha düşüktür.

**Boşanmış** insanların vadeli mevduat tekliflerine olumlu ve olumsuz cevap verme oranı en düşüktür.

- **Previous ve deposit:** Önceki kampanyalardan olumlu yanıt almış müşterilerin mevcut kampanyaya yanıt verip vermedikleri aşağıdaki grafikte görüntülenmektedir.

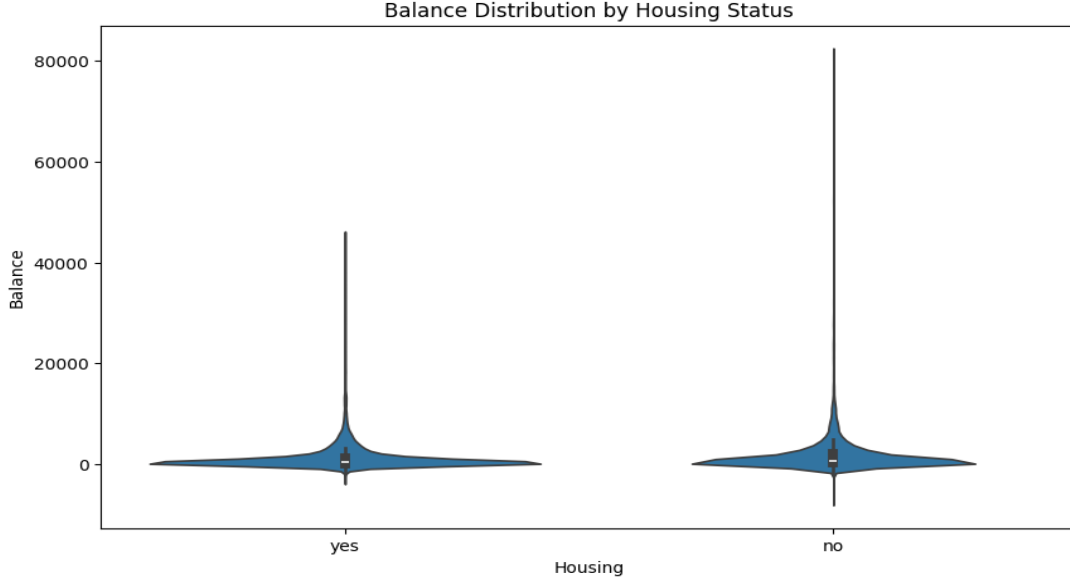


Daha önce hiçbir kampanyaya katılmayan müşteriler (0), mevcut kampanyayı kabul etme ve kabul etmeme oranı en yüksektir.



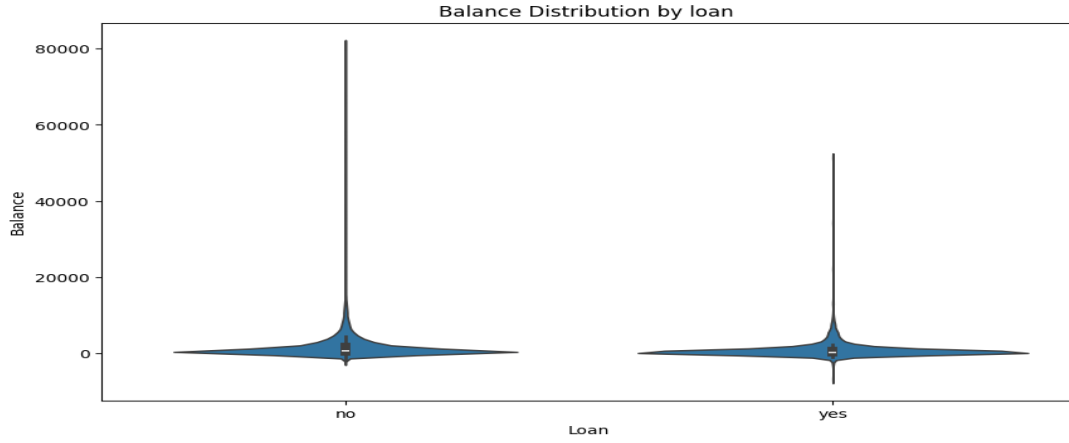
### Bankacılık İlişkisi ve Finansal Durum:

- **Housing ve balance:** Konut kredisi olan müşterilerin bakiyelerinin, konut kredisi olmayanlara göre farklı olup olmadığını grafiği aşağıda görüntülenmektedir.



**Konut kredisi olmayan** müşterilerin banka hesaplarındaki bakiye, **konut kredisi olan** müşterilere göre **daha yüksektir**.

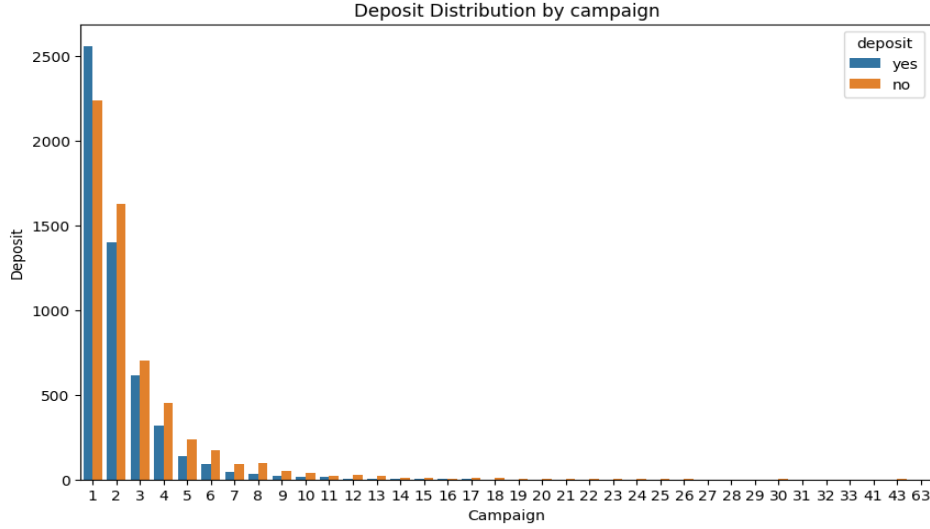
- **Loan ve balance:** Kişisel kredisi olan müşterilerin banka bakiyeleri üzerinde nasıl bir etkisi olduğunun grafiği aşağıda görüntülenmektedir.



**Kişisel kredisi olmayan** müşterilerin banka hesaplarındaki bakiye, **kişisel kredisi olan** müşterilere göre **daha yüksektir**.

## Kampanya Performansı:

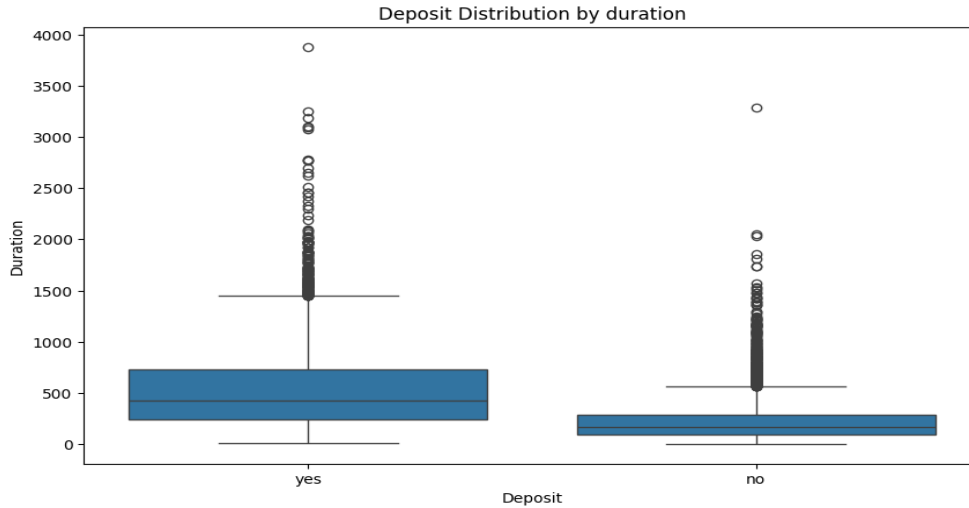
- **Campaign ve deposit:** Kampanya sırasında müşteriyle yapılan görüşme sayısının, müşterinin vadeli mevduat kabul etme olasılığı üzerindeki etkisinin grafiği aşağıda görüntülenmektedir.



Bir kez görüşme yapılan müşterilerde, vadeli mevduat tekliflerini kabul etme oranları en yüksektir.

Kampanya sırasında müşteriyle yapılan görüşme sayısının, müşterinin vadeli mevduat kabul etme olasılığında **olumlu etki yaratmamaktadır**.

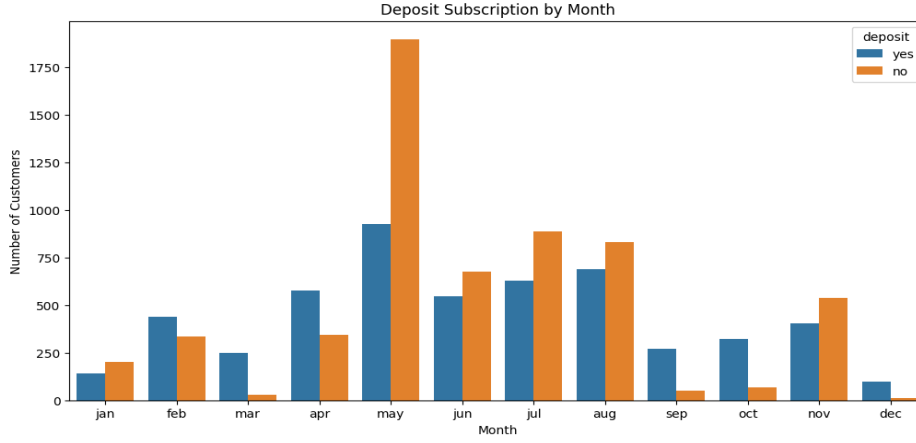
- **Duration ve deposit:** İletişim süresinin (telefon görüşmesi) kampanyanın başarılı olma olasılığı üzerindeki etkisinin grafiği aşağıda görüntülenmektedir.



İletişim süresi arttıkça, vadeli mevduat teklifinin kabul edilme olasılığının **arttığı** söylenebilir.

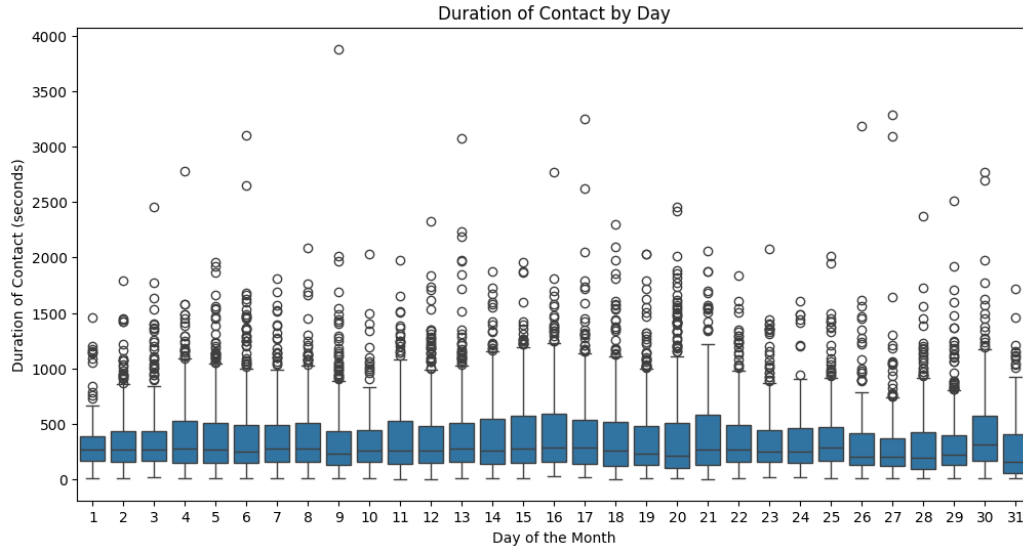
### Zaman Bazlı Analiz:

- **Month ve deposit:** Kampanyanın belirli aylarda daha başarılı olup olmadığını grafiği aşağıda görüntülenmektedir. Yılın farklı zamanlarında kampanyaların başarısı değişiyor mu?



Vadeli mevduat teklifinin yılın aylarına göre kabul edilme oranları arasında büyük bir fark yoktur. En yüksek kabul edilme oranı “Mayıs” ayında yaşanmıştır. En yüksek reddedilme oranı ise yine “Mayıs” ayında yaşanmıştır.

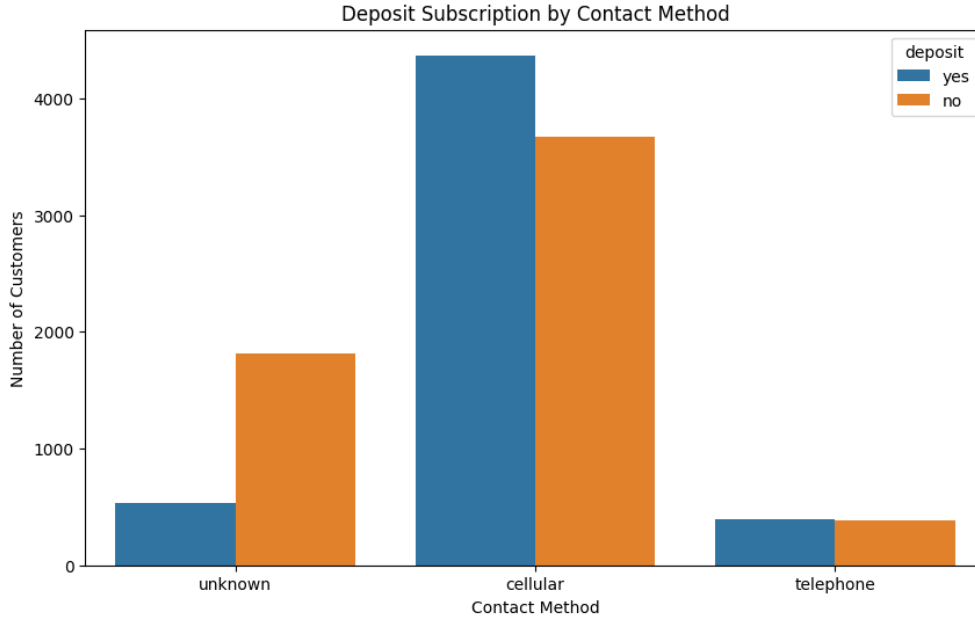
- **Day ve duration:** Görüşmenin yapıldığı günün ve görüşme süresinin müşteri yanıtları üzerindeki etkisinin grafiği aşağıda görüntülenmektedir.



Vadeli mevduat teklifinin ayın günlerine göre kabul edilme oranları arasında büyük bir fark yoktur.

### Müşteri İletişim Türü:

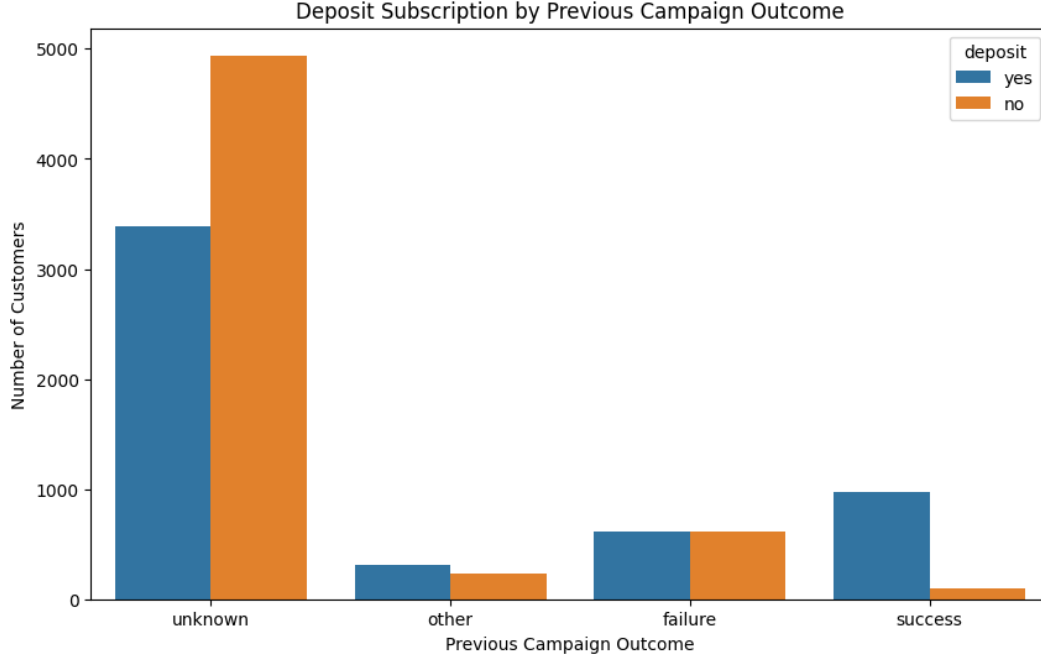
- **Contact** ve **deposit**: Farklı iletişim türlerinin (telefon, e-posta, vb.) kampanya başarısı üzerindeki etkisinin grafiği aşağıda görüntülenmektedir. Hangi iletişim türü daha etkili?



Cep telefonu ile iletişime geçilen müşterilerde, sabit hatlı telefonla ve bilinmeyen iletişim türü ile iletişime geçilen müşterilere göre vadeli mevduat kampanyasının daha başarılı olduğu söylenebilir.

### Önceki Kampanya Sonuçları ve Mevcut Durum:

- **Poutcome ve deposit:** Önceki kampanyaların sonuçlarının, mevcut kampanyadaki müşteri yanıtları üzerinde nasıl bir etkisi olduğunun grafiği aşağıda görüntülenmektedir. Önceki kampanyalarda olumlu yanıt verenler, mevcut kampanyaya nasıl yanıt veriyor?



Önceki kampanyalarda olumlu veya olumsuz yanıt verdikleri bilinmeyen müşterilerin, mevcut kampanyayı kabul etme oranları en yüksektir. Bu durumda önceki kampanyalardan harekete geçerek mevcut kampanya hakkında öngörü yapmak bu iki değişken arasında anlamlı olmayacaktır.

## Model Seçimi ve Eğitim

K-means algoritması ile müşteri segmentasyonu gerçekleştirilmiştir.

K-means algoritması için sadece gerekli sütunların seçilmiştir.

K-means modelini oluşturulup, eğitilmiştir.

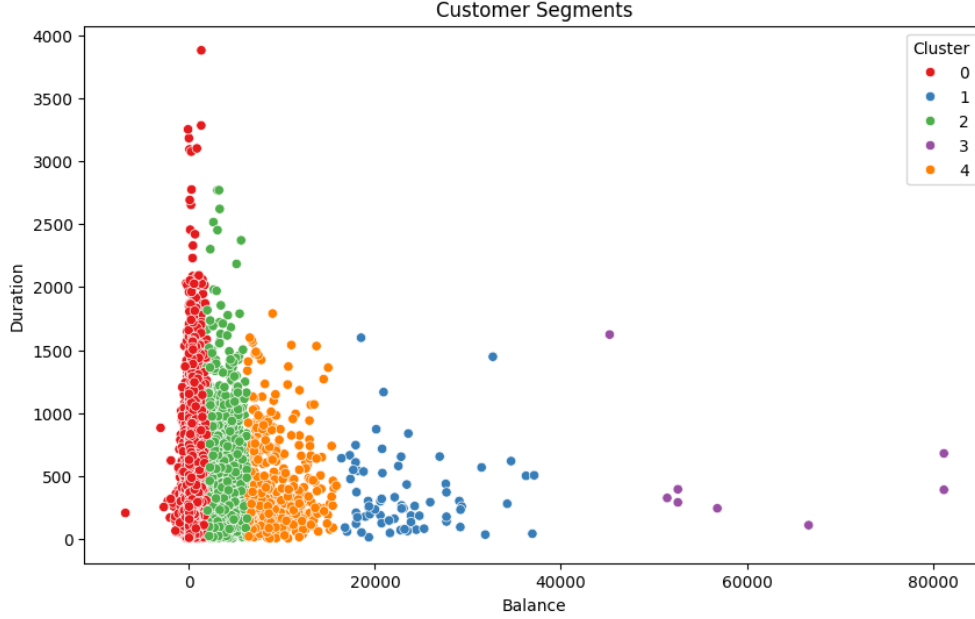
Müşteri segmentlerini veriye eklenip ve görüntülenmesi:

|   | age | job        | marital | education | default | balance | housing | loan | contact |
|---|-----|------------|---------|-----------|---------|---------|---------|------|---------|
| 0 | 59  | admin.     | married | secondary | no      | 2343    | yes     | no   | unknown |
| 1 | 56  | admin.     | married | secondary | no      | 45      | no      | no   | unknown |
| 2 | 41  | technician | married | secondary | no      | 1270    | yes     | no   | unknown |
| 3 | 55  | services   | married | secondary | no      | 2476    | yes     | no   | unknown |
| 4 | 54  | admin.     | married | tertiary  | no      | 184     | no      | no   | unknown |

|   | day | month | duration | campaign | pdays | previous | poutcome | deposit | Cluster |
|---|-----|-------|----------|----------|-------|----------|----------|---------|---------|
| 0 | 5   | may   | 1042     | 1        | -1    | 0        | unknown  | yes     | 2       |
| 1 | 5   | may   | 1467     | 1        | -1    | 0        | unknown  | yes     | 0       |
| 2 | 5   | may   | 1389     | 1        | -1    | 0        | unknown  | yes     | 0       |
| 3 | 5   | may   | 579      | 1        | -1    | 0        | unknown  | yes     | 2       |
| 4 | 5   | may   | 673      | 2        | -1    | 0        | unknown  | yes     | 0       |

## Sonuçların Yorumlanması ve Raporlama

Müşteri segmentlerinin görselleştirerek yorumlanması:



- **Küme 0 (Kırmızı):** Bu segment, çok düşük bakiyeye sahip ancak kampanya ile uzun süre iletişimde kalan müşterileri içeriyor. Bu durum, uzun süreli kampanya iletişiminin bu müşteriler üzerinde düşük bir finansal getiri sağladığını gösterebilir. Bu müşteriler, belki de daha farklı kampanya stratejileriyle daha etkili bir şekilde hedeflenebilir.
- **Küme 1 (Mavi):** Bu segment, kampanya iletişim süresi kısa olan (0-1000 arasında) ve genelde ortalama bakiyeye sahip müşterileri içeriyor. Bu müşteriler, kampanyaya uzun süre maruz kalmamış, ancak yine de makul bir bakiye tutuyorlar. Bu durum, kısa süreli iletişimle bile etkili bir sonuç alınabildiğini gösterebilir.
- **Küme 2 (Yeşil):** Bu segment, düşük bakiyeye sahip ve orta sürede kampanya iletişimi alan müşterileri içeriyor. Bu müşteriler, kampanyayla orta düzeyde etkileşimde bulunmuş, ancak finansal anlamda bankaya daha az katkı sağlamışlar.
- **Küme 3 (Mor):** Bu segment, çok yüksek bakiyeye sahip (40,000'den 80,000'e kadar) ve kısa ile orta sürede kampanya iletişimi alan müşterileri içeriyor. Bu müşteriler, kampanyaya çok uzun süre maruz kalmamış olmalarına rağmen bankaya yüksek finansal katkı sağlayan müşteriler olarak değerlendirilebilir. Banka için kritik ve yüksek değerli müşteriler olabilirler.

- **Küme 4 (Turuncu):** Bu segment, çok düşük bakiyeye sahip ve iletişim süresi düşük olan müşterilerden oluşuyor. Bu grup, kampanyaya en az maruz kalan ve bankaya en az katkı sağlayan müşteri grubu olarak değerlendirilebilir. Bu müşterilere yönelik kampanya stratejilerinin gözden geçirilmesi gerekebilir.

#### **Stratejik Öneriler:**

- **Küme 3 (Mor) Müşteriler:** Bu müşteriler, bankaya yüksek katkı sağlıyor ancak kampanya süresi çok uzun değil. Bu tür müşterilere daha fazla kişiselleştirilmiş ve değer katan kampanyalar sunarak sadakatleri artırılabilir.
- **Küme 0 (Kırmızı) Müşteriler:** Uzun süre kampanyaya maruz kalan ancak düşük katkı sağlayan bu müşteriler için daha etkili, belki de farklı türde kampanyalar düşünülmelidir.
- **Küme 4 (Turuncu) Müşteriler:** Bu müşterilere yönelik kampanya süresi artırılabilir ya da onlara özel yeni teklifler sunularak müşteri değerleri artırılabilir.



Müşteri segmentlerinin özet istatistikleri:

| cluster | age       | balance      | duration   |
|---------|-----------|--------------|------------|
| 0       | 40.381411 | 462.142144   | 366.850777 |
| 1       | 46.000000 | 23325.563380 | 356.126761 |
| 2       | 44.065541 | 3372.124628  | 390.979146 |
| 3       | 60.000000 | 60969.125000 | 506.625000 |
| 4       | 43.747228 | 9187.314856  | 385.598670 |

1. **Küme 0:**

- **Yaş:** Ortalama yaş 40.38 yıl.
- **Bakiye:** Ortalama bakiye 462 Euro, bu da oldukça düşük bir seviye.
- **Süre:** Kampanya ile ortalama 366.85 dakika boyunca iletişimde kalmışlar.
- **Yorum:** Bu segment, nispeten genç müşterilerden oluşuyor, bankada düşük bakiye tutuyorlar ve kampanya ile orta düzeyde bir süre etkileşimde bulunmuşlar. Bu müşteriler, bankaya çok yüksek bir finansal katkı sağlamıyor olabilir.

2. **Küme 1:**

- **Yaş:** Ortalama yaş 46 yıl.
- **Bakiye:** Ortalama bakiye 23,325 Euro, oldukça yüksek bir seviye.
- **Süre:** Kampanya ile ortalama 356.13 dakika boyunca iletişimde kalmışlar.
- **Yorum:** Bu segment, ortalama yaştan biraz üzerinde, bankaya yüksek bakiye ile önemli bir finansal katkı sağlıyor. Kampanya ile nispeten kısa bir süre etkileşimde bulunmuşlar. Bu müşteriler, banka için oldukça değerli olabilir ve onlara yönelik özel teklifler düşünülebilir.

3. **Küme 2:**

- **Yaş:** Ortalama yaş 44.07 yıl.
- **Bakiye:** Ortalama bakiye 3,372 Euro, düşük-orta bir seviye.
- **Süre:** Kampanya ile ortalama 390.98 dakika boyunca iletişimde kalmışlar.
- **Yorum:** Bu segment, orta yaş grubundaki müşterilerden oluşuyor, bankada düşük-orta seviyede bir bakiye tutuyorlar ve kampanya ile orta düzeyde bir süre etkileşimde bulunmuşlar. Bu müşteriler, kampanya ile biraz daha fazla süre etkileşimde bulunmuş ancak hala düşük bir finansal katkı sağlıyor olabilir.

#### 4. Küme 3:

- **Yaş:** Ortalama yaş 60 yıl.
- **Bakiye:** Ortalama bakiye 60,969 Euro, çok yüksek bir seviye.
- **Süre:** Kampanya ile ortalama 506.63 dakika boyunca iletişimde kalmışlar.
- **Yorum:** Bu segment, en yaşlı müşteri grubunu içeriyor, bankada en yüksek bakiye ile çok önemli bir finansal katkı sağlıyorlar. Ayrıca, kampanya ile uzun süre etkileşimde bulunmuşlar. Bu müşteriler, bankanın en değerli müşterileri arasında yer alabilir ve sadakat programları gibi stratejilerle özel olarak hedeflenebilirler.

#### 5. Küme 4:

- **Yaş:** Ortalama yaş 43.75 yıl.
- **Bakiye:** Ortalama bakiye 9,187 Euro, orta seviye.
- **Süre:** Kampanya ile ortalama 385.60 dakika boyunca iletişimde kalmışlar.
- **Yorum:** Bu segment, orta yaş grubundaki müşterilerden oluşuyor, bankada orta seviyede bir bakiye tutuyorlar ve kampanya ile orta düzeyde bir süre etkileşimde bulunmuşlar. Bu müşteriler, bankaya ortalama bir finansal katkı sağlıyor olabilir ve onların sadakatlerini artırmak için ek teşvikler değerlendirilebilir.

#### Genel Sonuç:

- **Küme 1 ve 3:** Bu kümeler, bankaya en yüksek finansal katkıyı sağlıyor. Küme 3 özellikle dikkat çekici çünkü hem yaşlı hem de uzun süre kampanya ile etkileşimde bulunmuş. Bu müşteriler için özel stratejiler geliştirilebilir.
- **Küme 0 ve 2:** Bu kümeler, düşük bakiyelere sahip, ancak özellikle Küme 2 kampanya ile daha uzun süre etkileşimde kalmış. Onların finansal katkısını artırmak için farklı kampanya stratejileri değerlendirilebilir.
- **Küme 4:** Orta düzeyde bir profil sunuyor; bankanın bu segmenti nasıl daha üst seviyelere çekebileceğini düşünmesi gerekebilir.

## Kaynak:

Veri: <https://www.kaggle.com/code/aleksandradeis/bank-marketing-analysis/input>

Python Görselleştirmeler ve Analizler: [https://colab.research.google.com/drive/1B\\_zpAw504uuQqKWK-fvJoF-nl3637aJq?usp=sharing](https://colab.research.google.com/drive/1B_zpAw504uuQqKWK-fvJoF-nl3637aJq?usp=sharing)