

Didem sunum script

Slide 1: Introduction

Title: Introduction to Heart Failure Clinical Records Dataset

Script:

"Welcome everyone. Today, I will be presenting our analysis on the heart failure clinical records dataset. This dataset contains clinical information from 299 heart failure patients and aims to identify prognostic factors for survival and risk indicators associated with heart failure."

Slide 2: Data Overview

Title: Brief Overview of Key Variables

Script:

"Let's start by introducing our key variables:

- **Age:** Patient's age at the time of follow-up.
- **Anaemia:** Binary variable indicating presence of anemia.
- **CPK_Level:** Measures the enzyme levels related to heart stress.
- **Diabetes:** Indicates if the patient has diabetes.
- **EjectionFraction:** The percentage of blood leaving the heart each beat.
- **High Blood Pressure:** Indicates if the patient has hypertension.
- **Platelets:** Measures blood platelet count.
- **Serum Creatinine:** Indicates kidney function.
- **Serum Sodium:** Indicates sodium level in the blood.
- **Gender:** Patient's gender.
- **Smoking:** Indicates if the patient is a smoker.

- **Follow-Up Days:** Duration of follow-up in days.
 - **Death Event:** Indicates if the patient died during the follow-up period."
-

Slide 3: Data Cleaning and Formatting

Title: Data Cleaning and Formatting

Script:

"We cleaned and formatted the data to prepare it for analysis. This involved:

- **Renaming Variables:** For better clarity.
 - **Creating New Variables:** Like AgeGroup and HighRisk.
 - **Filtering the Data:** To remove extreme outliers."
-

Slide 4: Checking Model Assumptions

Title: Checking Model Assumptions

Script:

"We checked key assumptions to ensure the validity of our model:

- **Normality:** Assessed using Q-Q plots.
 - **Linearity:** Ensured that the relationship between predictors and the response is linear.
 - **Homoscedasticity:** Checked for constant variance of residuals."
-

Slide 5: Model Choice

Title: Model Choice

Script:

"For our analysis, we chose multiple linear regression. This model helps us:

- Quantify relationships between variables and death events.
 - Identify significant predictors.
 - Guide clinical decisions to forecast patient outcomes."
-

Slide 6: Packages and Tools

Title: Packages and Tools

Script:

"We used the following R packages for our analysis:

- **dplyr**: For data manipulation and cleaning.
 - **readr**: For reading CSV files.
 - **ggplot2**: For data visualization.
 - **stats (Base R)**: For statistical modeling."
-

Slide 7: Descriptive Statistics

Title: Descriptive Statistics

Script:

"Here are the summary statistics of our variables:

- **Age**: Mean 60.83, Median 60, Std. Dev. 11.89, Min 40, Max 95.
- **Anaemia**: Mean 0.43, Median 0, Std. Dev. 0.50, Min 0, Max 1.
- **CPK_Level**: Mean 581.84, Median 250, Std. Dev. 970.29, Min 23, Max 7861.
- **Diabetes**: Mean 0.42, Median 0, Std. Dev. 0.49, Min 0, Max 1.
- **EjectionFraction**: Mean 38.08, Median 38, Std. Dev. 11.83, Min 14, Max 80.
- **HighBloodPressure**: Mean 0.35, Median 0, Std. Dev. 0.48, Min 0, Max 1.
- **Platelets**: Mean 263358.03, Median 262000, Std. Dev. 97804.24, Min 25100, Max 850000.
- **SerumCreatinine**: Mean 1.39, Median 1.1, Std. Dev. 1.03, Min 0.5, Max 9.4.
- **SerumSodium**: Mean 136.63, Median 137, Std. Dev. 4.41, Min 113, Max 148.
- **Gender**: Mean 0.65, Median 1, Std. Dev. 0.48, Min 0, Max 1.
- **Smoker**: Mean 0.32, Median 0, Std. Dev. 0.47, Min 0, Max 1.
- **FollowUpDays**: Mean 130.26, Median 115, Std. Dev. 77.57, Min 4, Max 285.
- **DeathEvent**: Mean 0.32, Median 0, Std. Dev. 0.47, Min 0, Max 1."

Slide 8: Regression Model Summary

Title: Regression Model Summary

Script:

"Our regression model summary highlights:

- **Intercept:** -0.8522.
 - **Significant Predictors:** CPK_Level, EjectionFraction, and SerumCreatinine.
 - **Model Performance:** R-squared 0.455, Adjusted R-squared 0.435."
-

Slide 9: Key Findings

Title: Key Findings

Script:

"Our key findings are:

- Higher levels of CPK_Level increase the risk of death.
 - Lower ejection fraction increases the risk of death.
 - Higher serum creatinine levels increase the risk of death.
 - The model explains approximately 45.5% of the variance in death events."
-

Slide 10: Visual Analysis of Key Health Indicators

Title: Visual Analysis of Key Health Indicators

Script:

"Let's look at some visualizations to better understand our data:

- **Age Distribution:** Shows the age range of patients.
- **Frequency of Anaemia:** Displays the count of patients with and without anaemia.
- **Serum Creatinine Levels:** A histogram showing the distribution of serum creatinine levels.

- **Ejection Fraction by Death Event:** A box plot showing the distribution of ejection fraction by death event."
-

Slide 11: Conclusion

Title: Conclusion

Script:

"In conclusion:

- The regression analysis indicates that CPK_Level, EjectionFraction, and SerumCreatinine are significant predictors of DeathEvent.
 - Our model helps identify high-risk patients and improve treatment strategies.
 - These insights can guide clinicians in making informed decisions to improve patient outcomes."
-

This script provides a clear, concise, and logical flow for presenting your analysis. Make sure to adjust the script based on the specific content and visualizations in your final PPTX presentation.

Slayt 1: Giriş

Başlık: Kalp Yetmezliği Klinik Kayıtlar Veri Setine Giriş

Metin:

"Herkese hoş geldiniz. Bugün, kalp yetmezliği klinik kayıtları veri seti üzerindeki analizimizi sunacağım. Bu veri seti, 299 kalp yetmezliği hastasından klinik bilgileri içermekte ve hayatta kalma için prognostik faktörleri ve kalp yetmezliği ile ilişkili risk göstergelerini belirlemeyi amaçlamaktadır."

Slayt 2: Veri Genel Bakışı

Başlık: Anahtar Değişkenlere Kısa Genel Bakış

Metin:

"Anahtar değişkenlerimizi tanıtarak başlayalım:

- **Yaş:** Takip sırasındaki hastanın yaşı.
- **Anemi:** Anemi varlığını belirten ikili değişken.
- **CPK_Düzeyi:** Kalp kası stresine bağlı enzim seviyelerini ölçer.
- **Diyabet:** Hastanın diyabet durumunu belirtir.
- **Atım Hacmi:** Kalbin her atımda pompaladığı kan yüzdesi.
- **Yüksek Tansiyon:** Hastanın hipertansiyon durumunu belirtir.
- **Trombositler:** Kan trombosit sayısını ölçer.
- **Serum Kreatinin:** Böbrek fonksiyonunu gösterir.
- **Serum Sodyum:** Kanda sodyum seviyesini gösterir.
- **Cinsiyet:** Hastanın cinsiyeti.
- **Sigara:** Hastanın sigara içip içmediğini belirtir.
- **Takip Günleri:** Takip süresi (gün).
- **Ölüm Olayı:** Takip süresi içinde hastanın ölüp ölmediğini belirtir."

Slayt 3: Veri Temizleme ve Biçimlendirme

Başlık: Veri Temizleme ve Biçimlendirme

Metin:

"Verileri analiz için hazırlamak amacıyla temizledik ve biçimlendirdik. Bu işlem şunları içerir:

- **Değişken Adlarını Yeniden Adlandırma:** Daha iyi anlaşılabilirlik için.
- **Yeni Değişkenler Oluşturma:** YaşGrubu ve YüksekRisk gibi.
- **Veriyi Filtreleme:** Aşırı uç değerleri çıkarmak için."

Slayt 4: Model Varsayımlarını Kontrol Etme

Başlık: Model Varsayımlarını Kontrol Etme

Metin:

"Modelimizin geçerliliğini sağlamak için temel varsayımları kontrol ettik:

- **Normallik:** Q-Q grafiklerini kullanarak değerlendirdik.
 - **Doğrusallık:** Bağımsız değişkenler ve bağımlı değişken arasındaki ilişkinin doğrusal olduğunu sağladık.
 - **Homoscedasticity:** Artıkların sabit varyansını kontrol ettik."
-

Slayt 5: Model Seçimi

Başlık: Model Seçimi**Metin:**

"Analizimiz için çoklu doğrusal regresyon modelini seçtik. Bu model bize şunları sağlar:

- Değişkenler ve ölüm olayları arasındaki ilişkileri nicel olarak belirlemek.
 - Önemli tahmin edicileri tanımlamak.
 - Klinik kararları yönlendirmek için hasta sonuçlarını tahmin eden bir destek aracı sağlamak."
-

Slayt 6: Paketler ve Araçlar

Başlık: Paketler ve Araçlar**Metin:**

"Analizimiz için aşağıdaki R paketlerini kullandık:

- **dplyr:** Veri manipülasyonu ve temizleme için.
 - **readr:** CSV dosyalarını okumak için.
 - **ggplot2:** Veri görselleştirme için.
 - **stats (Base R):** İstatistiksel modelleme için."
-

Slayt 7: Betimleyici İstatistikler

Başlık: Betimleyici İstatistikler

Metin:

"İşte değişkenlerimizin özet istatistikleri:

- **Yaş:** Ort. 60.83, Ortanca 60, Std. Sapma 11.89, Min 40, Maks 95.
- **Anemi:** Ort. 0.43, Ortanca 0, Std. Sapma 0.50, Min 0, Maks 1.
- **CPK_Düzeyi:** Ort. 581.84, Ortanca 250, Std. Sapma 970.29, Min 23, Maks 7861.
- **Diyabet:** Ort. 0.42, Ortanca 0, Std. Sapma 0.49, Min 0, Maks 1.
- **Atım Hacmi:** Ort. 38.08, Ortanca 38, Std. Sapma 11.83, Min 14, Maks 80.
- **Yüksek Tansiyon:** Ort. 0.35, Ortanca 0, Std. Sapma 0.48, Min 0, Maks 1.
- **Trombositler:** Ort. 263358.03, Ortanca 262000, Std. Sapma 97804.24, Min 25100, Maks 850000.
- **Serum Kreatinin:** Ort. 1.39, Ortanca 1.1, Std. Sapma 1.03, Min 0.5, Maks 9.4.
- **Serum Sodyum:** Ort. 136.63, Ortanca 137, Std. Sapma 4.41, Min 113, Maks 148.
- **Cinsiyet:** Ort. 0.65, Ortanca 1, Std. Sapma 0.48, Min 0, Maks 1.
- **Sigara:** Ort. 0.32, Ortanca 0, Std. Sapma 0.47, Min 0, Maks 1.
- **Takip Günleri:** Ort. 130.26, Ortanca 115, Std. Sapma 77.57, Min 4, Maks 285.
- **Ölüm Olayı:** Ort. 0.32, Ortanca 0, Std. Sapma 0.47, Min 0, Maks 1."

Slayt 8: Regresyon Modeli Özeti

Başlık: Regresyon Modeli Özeti**Metin:**

"Regresyon modelimizden elde edilen özet sonuçlar şunlardır:

- **Kesme Noktası:** -0.8522.
- **Önemli Tahmin Ediciler:** CPK_Düzeyi, Atım Hacmi ve Serum Kreatinin.
- **Model Performansı:** R-kare 0.455, Ayarlanmış R-kare 0.435."

Slayt 9: Ana Bulgular

Başlık: Ana Bulgular

Metin:

"Ana bulgularımız şunlardır:

- Yüksek CPK_Düzeyi ölüm riskini artırır.
 - Düşük atım hacmi ölüm riskini artırır.
 - Yüksek serum kreatinin seviyeleri ölüm riskini artırır.
 - Model, ölüm olaylarındaki varyansın yaklaşık %45.5'ini açıklar."
-

Slayt 10: Anahtar Sağlık Göstergelerinin Görsel Analizi

Başlık: Anahtar Sağlık Göstergelerinin Görsel Analizi

Metin:

"Verimizi daha iyi anlamak için bazı görselleştirmelere bakalım:

- **Yaş Dağılımı:** Hastaların yaş aralığını gösterir.
 - **Anemi Sıklığı:** Anemili ve anemisiz hasta sayısını gösterir.
 - **Serum Kreatinin Seviyeleri:** Serum kreatinin seviyelerinin dağılımını gösteren histogram.
 - **Ölüm Olayına Göre Atım Hacmi:** Ölüm olayına göre atım hacmi dağılımını gösteren kutu grafiği."
-

Slayt 11: Sonuç

Başlık: Sonuç

Metin:

"Sonuç olarak:

- Regresyon analizi, CPK_Düzeyi, Atım Hacmi ve Serum Kreatinin'in Ölüm Olayı için önemli tahmin ediciler olduğunu gösterir.
- Modelimiz, yüksek riskli hastaları tanımlamaya ve tedavi stratejilerini iyileştirmeye yardımcı olur.

- Bu bulgular, klinisyenlerin hasta sonuçlarını iyileştirmek için bilinçli kararlar vermesine rehberlik edebilir."

Bu metin, sunumunuzun her slaytı için net, öz ve mantıklı bir akış sağlar. Sunumun içeriğine ve görselleştirmelere bağlı olarak metni ayarladığınızdan emin olun.