

Analisis Exploratorio Kaggle Dataset

Jeniffer_Funez

2023-11-22

```
# Cargar los datos
data <- read.csv("C:/Users/jenif/OneDrive/Documentos/Proyecto_EDA_&_Analisis_exploratorio/Ghibli-charac

# Esto asegura la edad numerica
data$age <- as.numeric(data$age)
```

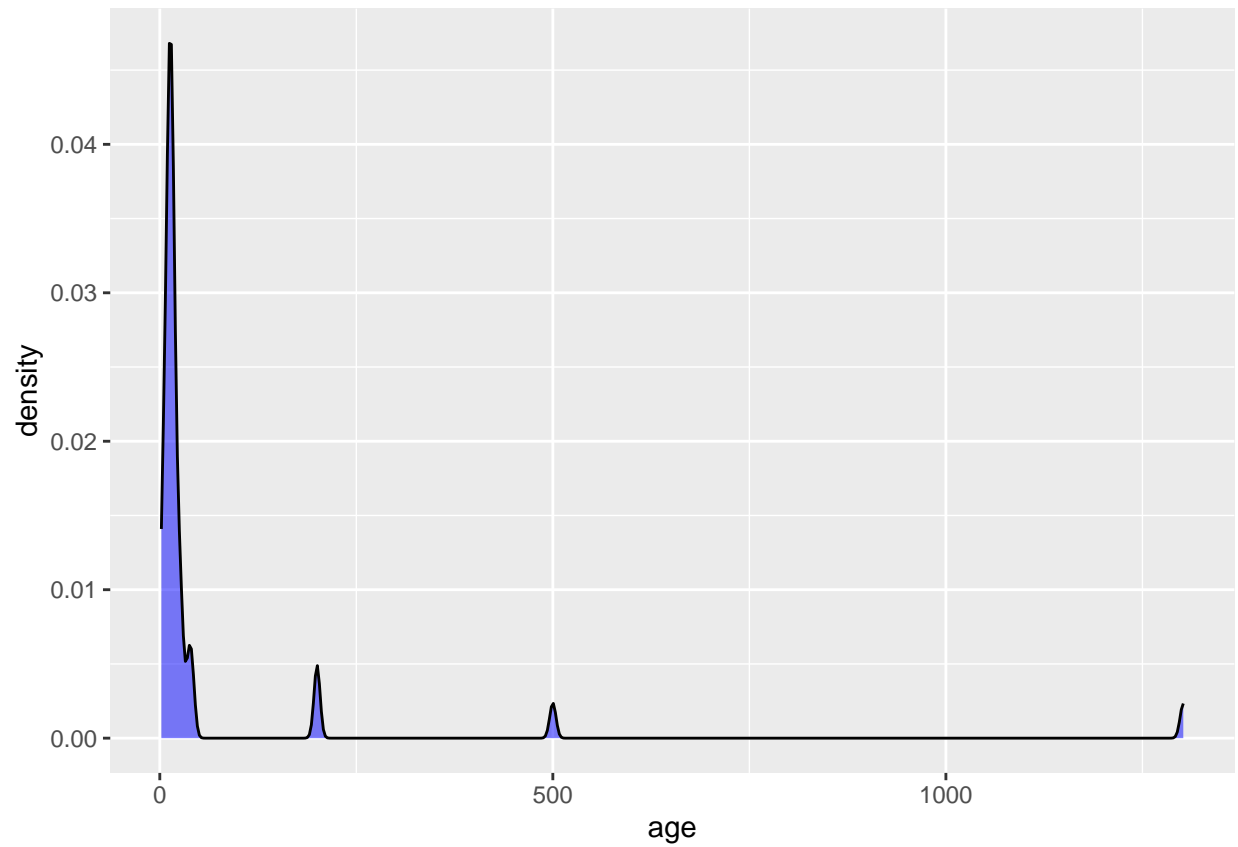
```
## Warning: NAs introducidos por coerción
```

```
# Limpio y preparo los datos para la grafica
clean_data <- data %>%
  drop_na(age) # esto elimina los valores NA para la columna de edad si existen

# Creando grafico de densidad para la edad
age_density <- ggplot(clean_data, aes(x=age)) +
  geom_density(fill="blue", alpha=0.5) #color
  labs(title="Densidad de Edad de los Personajes Ghibli", x="Edad", y="Densidad") +
  theme_minimal()# añadi un tema minimalista
```

```
## NULL
```

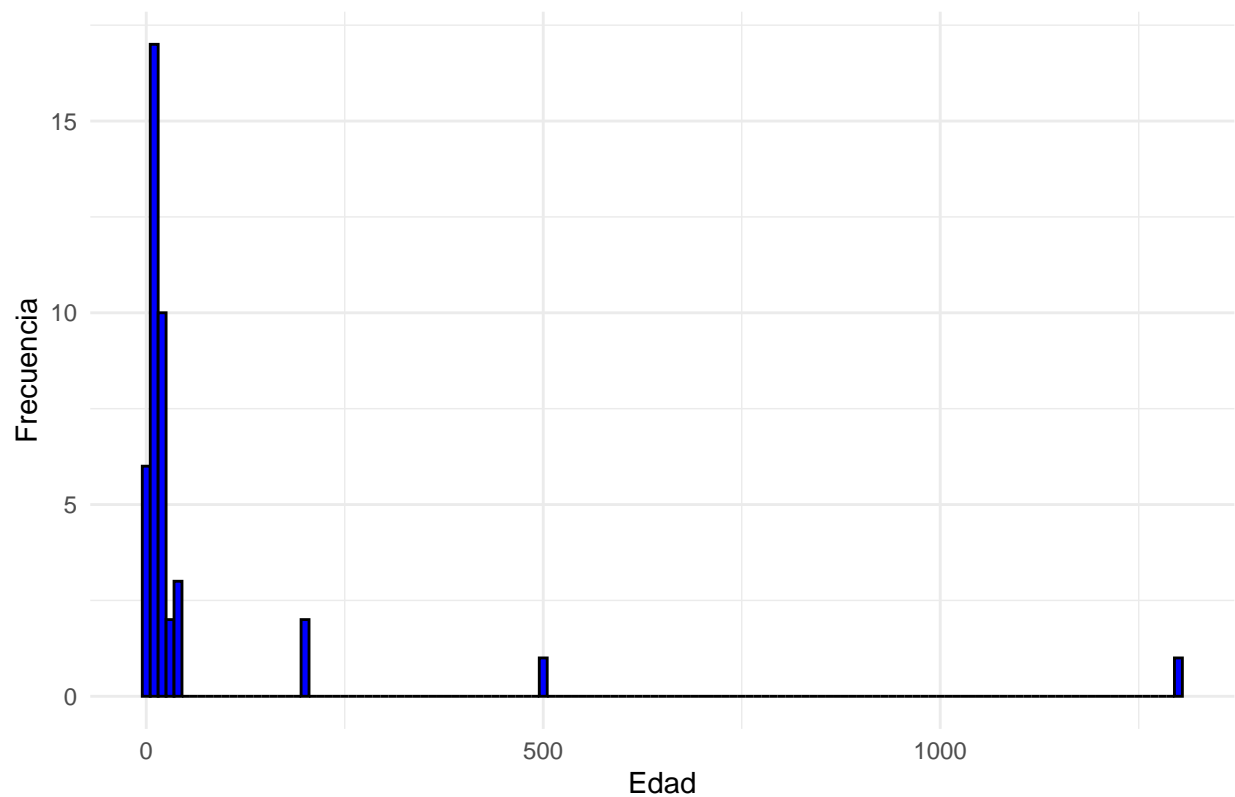
```
age_density
```



```
# creando el histograma para la edad
age_histogram <- ggplot(clean_data, aes(x=age)) +
  geom_histogram(binwidth=10, fill="blue", color="black") +
  labs(title="Histograma de Edades de los Personajes de Ghibli", x="Edad", y="Frecuencia") +
  theme_minimal()

age_histogram
```

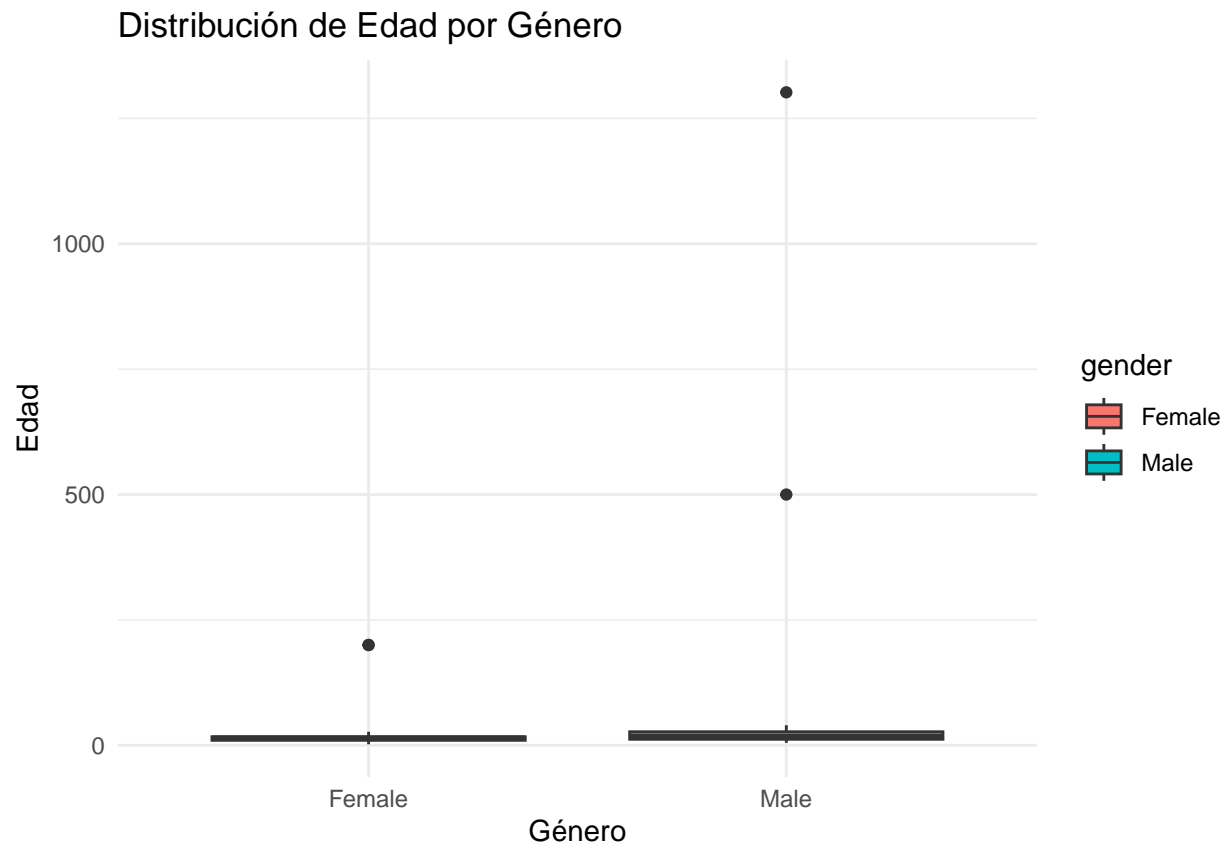
Histograma de Edades de los Personajes de Ghibli



```
# Boxplot para la edad por género
boxplot_age_gender <- ggplot(data, aes(x=gender, y=age, fill=gender)) +
  geom_boxplot() +
  labs(title="Distribución de Edad por Género", x="Género", y="Edad") +
  theme_minimal()

boxplot_age_gender
```

```
## Warning: Removed 14 rows containing non-finite values ('stat_boxplot()').
```

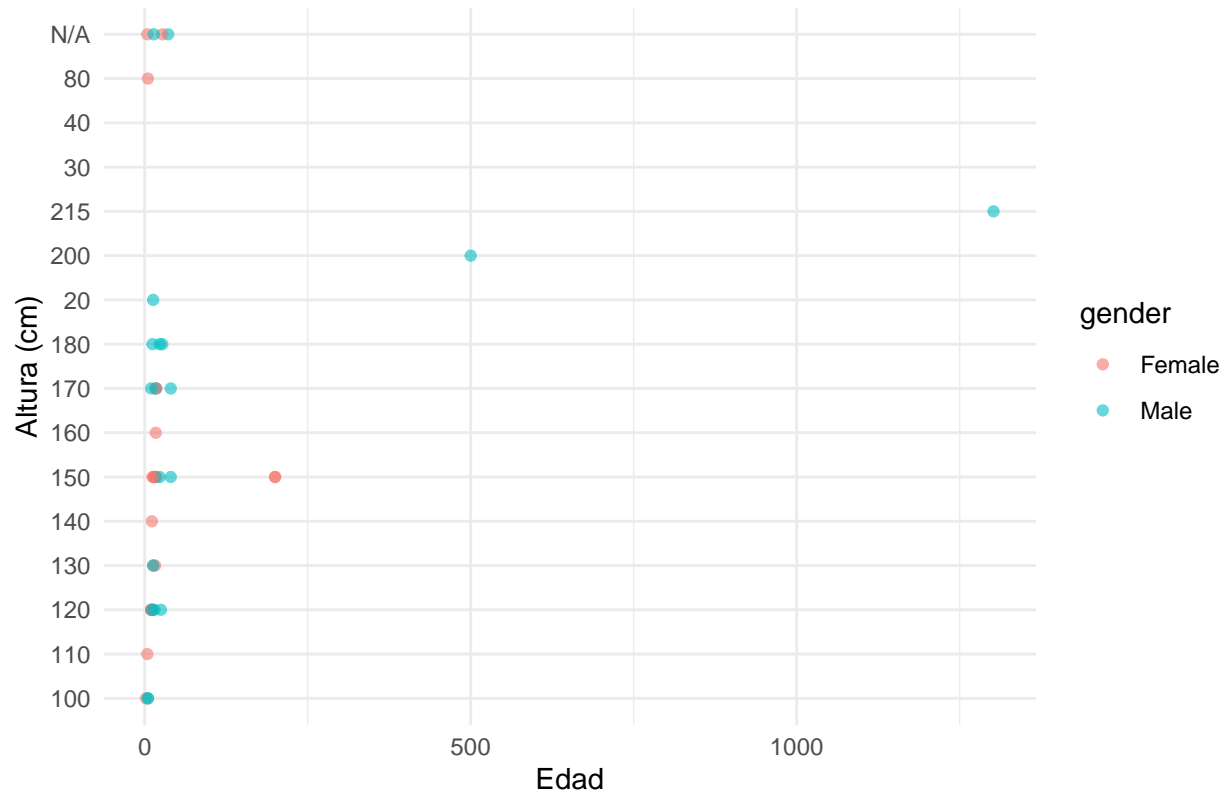


```
# Scatterplot para la relación entre edad y altura
scatterplot_age_height <- ggplot(data, aes(x=age, y=height..cm.)) +
  geom_point(aes(color=gender), alpha=0.6) +
  labs(title="Relación entre Edad y Altura", x="Edad", y="Altura (cm)") +
  theme_minimal()

scatterplot_age_height
```

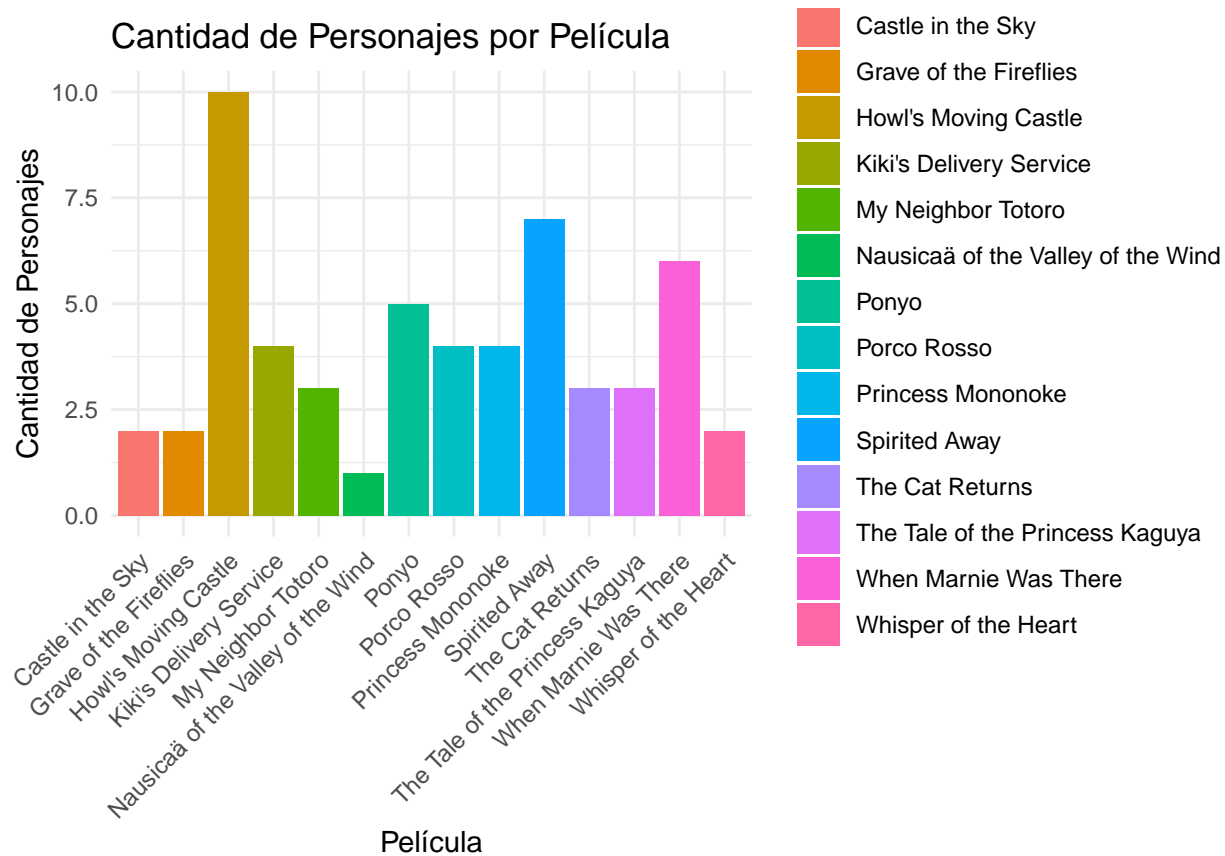
```
## Warning: Removed 14 rows containing missing values ('geom_point()').
```

Relación entre Edad y Altura



```
# Barplot para contar personajes por película
barplot_characters_movie <- ggplot(data, aes(x=movie)) +
  geom_bar(aes(fill=movie)) +
  theme_minimal() +
  labs(title="Cantidad de Personajes por Película", x="Película", y="Cantidad de Personajes") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))

barplot_characters_movie
```



```
# Lineplot para la tendencia de altura a lo largo del tiempo
lineplot_height_release <- ggplot(data, aes(x=release.date, y=height..cm., group=1)) +
  geom_line() +
  labs(title="Tendencia de Altura de los Personajes a lo largo del Tiempo", x="Año de Lanzamiento", y="")
theme_minimal()

lineplot_height_release
```

Tendencia de Altura de los Personajes a lo largo del Tiempo

