Traffic Data Analysis

WITH PYTHON

Pei Yii Ng (Heather)

8 May, 2024

codecademy

BACKGROUND

In recent years, the automotive industry has witnessed a remarkable surge in **integrating advanced safety features into vehicles**. From traditional safety measures like seatbelts and airbags to cutting-edge technologies such as sensors, warning signals, and automated braking systems, manufacturers have continually aimed to enhance vehicle safety standards.

However, this progression has occurred alongside the proliferation of **distracting technologies within cars**, including GPS navigation, streaming services, and smartphone connectivity. Consequently, a pertinent question arises: amidst these technological advancements, **has the increasing use of smartphones resulted in deteriorating driving behaviors?**

ANALYSIS OBJECTIVES

VISUALISE

trends in car crashes over time

Create visual representations of car crash data to identify temporal patterns and trends in traffic safety outcomes.

PLOT

changes in key variables

Analyze data on smartphone usage among drivers and visualize changes over time, exploring potential associations with collision rates.

CORRELATION

of each variable with collision rate

Conduct statistical analysis to determine the strength and direction of the relationship between smartphone usage and collision rates, informing our understanding of driving behaviors.

PREDICT

collision rates with linear regression

Utilize predictive modeling techniques, such as linear regression, to forecast collision rates based on smartphone usage and other relevant factors, facilitating evidence-based policymaking and intervention strategies.

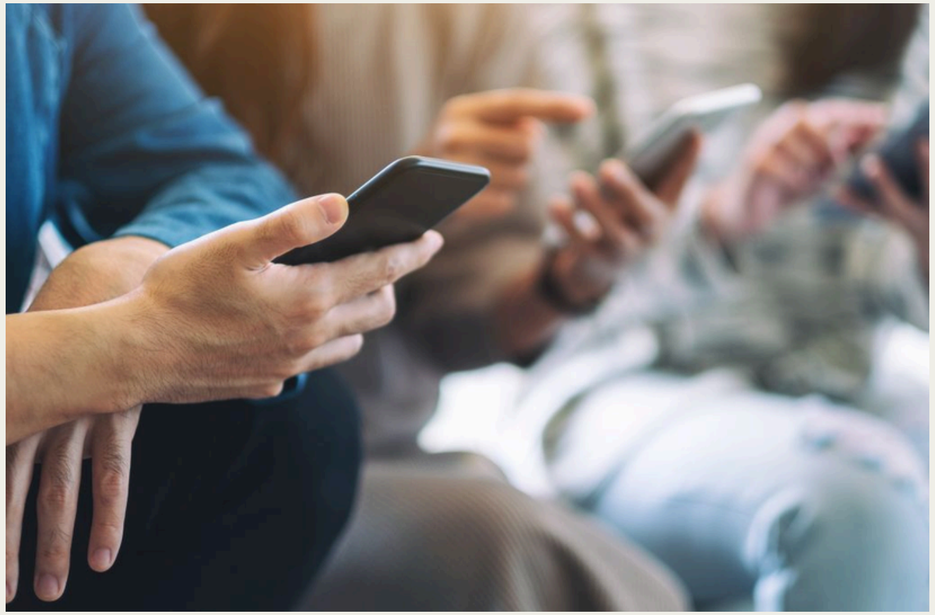
DATA SOURCE #1

Collisions Dataset Source

Monthly total numbers of collisions in the United States from 2006 to 2020, normalized to the population size.

Date	Crashes_Per_100k	Season
2006-01-01	**169.176541**	Winter
2006-02-01	**154.028836**	Winter
2006-03-01	**159.930002**	Spring
2006-04-01	**155.741270**	Spring
2006-05-01	**168.179208**	Spring

First 5 rows of the dataset



DATA SOURCE #2

Smartphone Usage Dataset Source

Quarterly percentage of American adults reporting owning and using a smartphone, from 2011 to 2019.

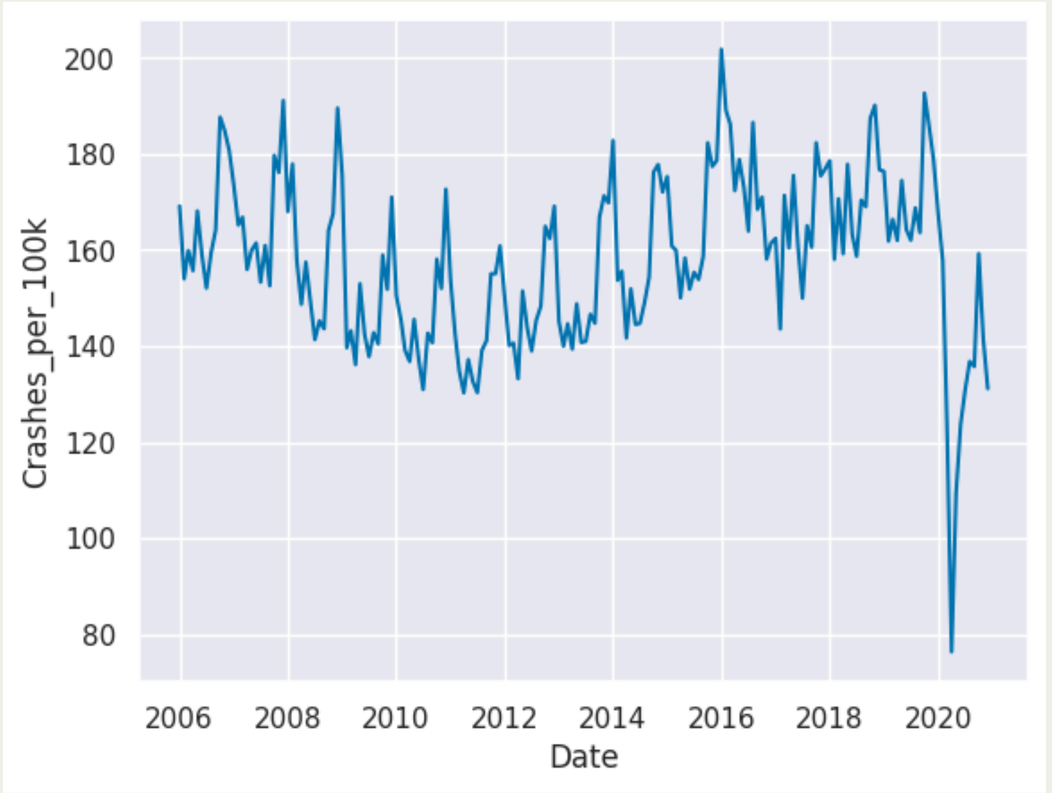
Month_Year	Crashes_Per_100k	Season	Smartphone_Survey_Date	Smartphone_usage
Apr-12	**133.213685**	Spring	4/3/12	46
Apr-15	**150.077792**	Spring	4/12/15	67
Apr-16	**172.401948**	Spring	4/4/16	72
Aug-12	**145.403147**	Summer	8/5/12	44
Dec-12	**169.160811**	Winter	12/9/12	45

First 5 rows of the dataset

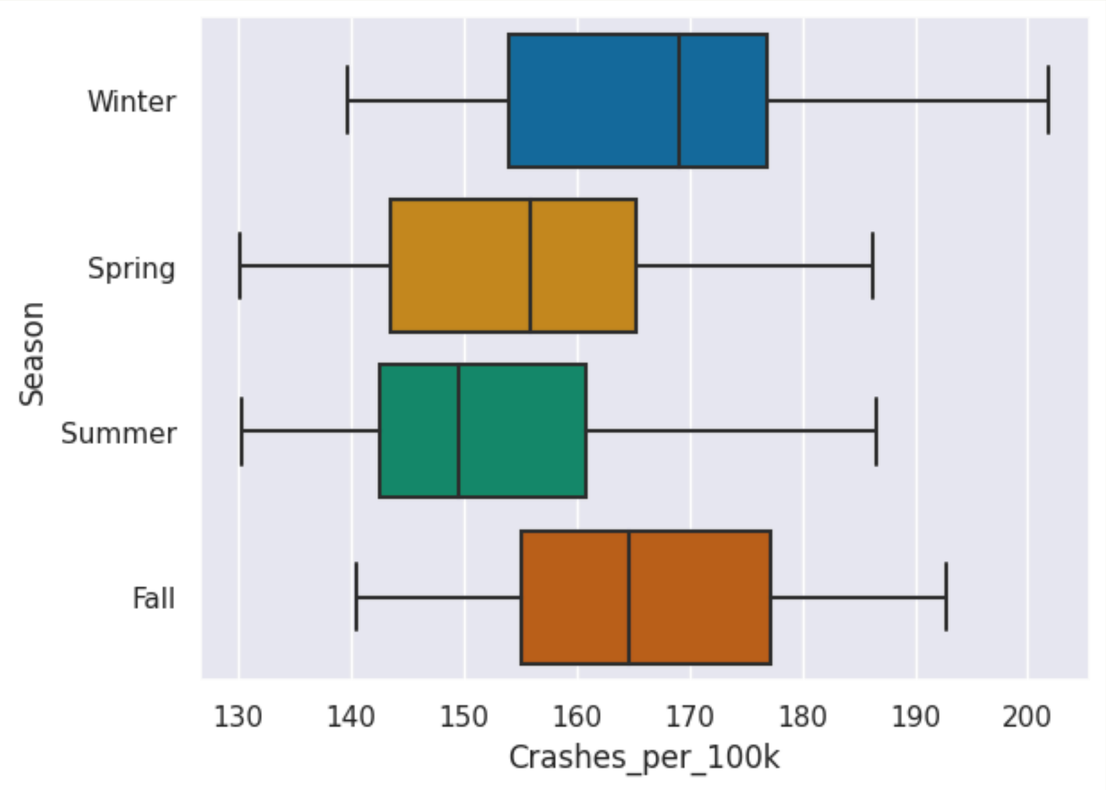
VISUALISE TRAFFIC SAFETY DATA

Visual inspection of the collision data revealed a decreasing trend in crash rates from 2006 to around 2010, followed by an increasing trend thereafter. There appear to be cyclical patterns, which might imply that crash rates vary by season.

The crash rates for 2020 showed significant deviation from previous years, likely due to external factors such as the COVID-19 pandemic.



SEASONAL VARIATIONS



Analysis of seasonal variations indicated higher crash rates during winter and fall compared to spring and summer.

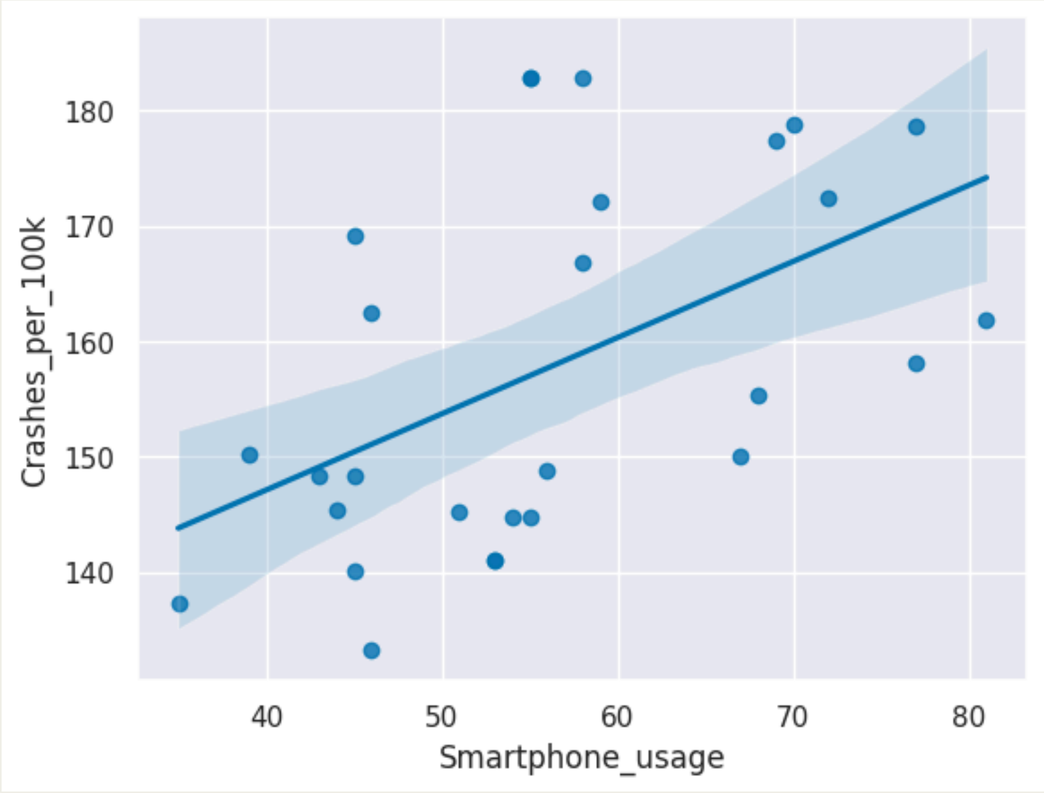
This suggests that environmental factors may influence driving conditions and accident likelihood.

SMARTPHONE USAGE V/S CRASH RATES

PEARSON'S R = 0.513

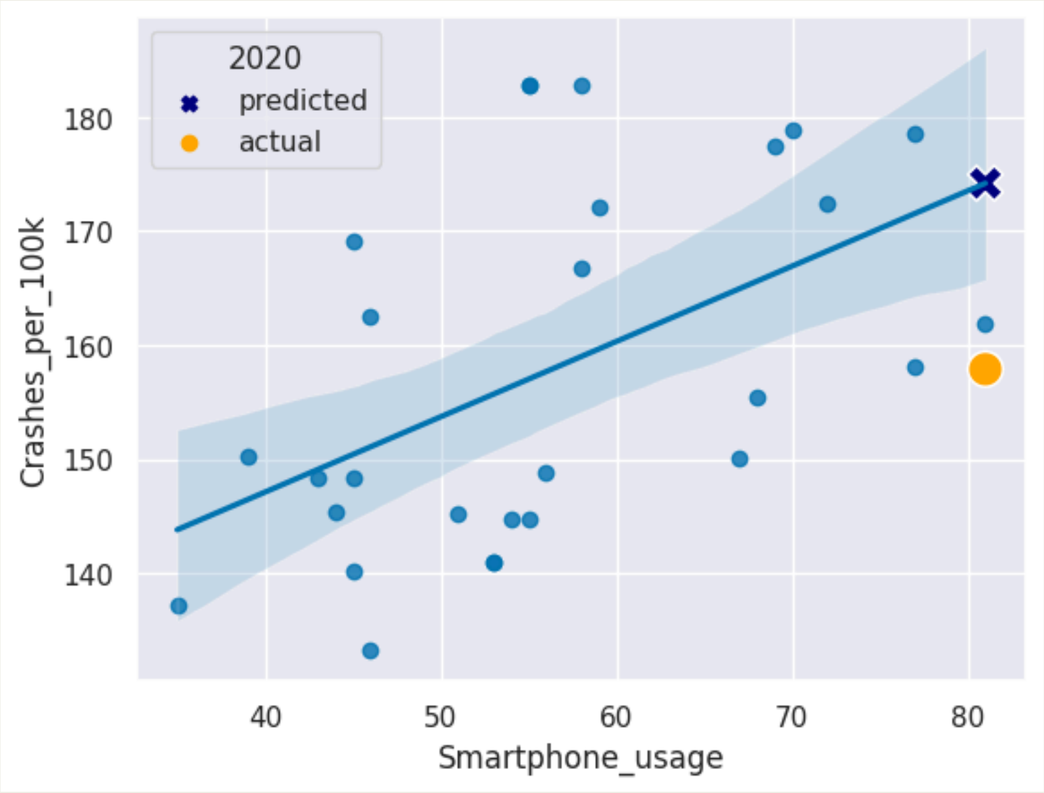
P = 0.005

Scatter plot analysis and Pearson correlation coefficient calculation revealed a moderately strong positive correlation between smartphone usage rates and crash rates. This suggests that higher smartphone usage is associated with increased crash rates.



LINEAR REGRESSION

```
Crashes_per_100k = 120.6637 + (0.6610 * Smartphone_usage)
```



The regression equation **predicted a crash rate of 174.21** per 100,000 people for 2020, assuming smartphone usage remained consistent with 2019 levels. However, the **actual crash rate** for February 2020, obtained from traffic safety data, was **lower at 157.89** per 100,000 people.

IMPLICATIONS

- Urgent need for awareness campaigns and policy interventions to address distracted driving and promote safer driving behaviours.
- Continued monitoring of smartphone usage trends essential for informed policymaking and targeted intervention strategies.