



# IOWA LIQUOR SALES

---

1st Jan. - 31st Dec. 2022

Data Science

Kai  
Schillinger

Luis  
Rastetter

## Data Selection

Zur Vorbereitung auf das Projekt forschten wir im Internet auf diversen Seiten und diskutiert, auf was sich unsere Ausarbeitung inhaltlich beziehen soll.

Da wir in unserem Leben beide schon positive als auch negative Erfahrungen im Umgang mit Alkohol im privaten Umfeld erfahren, waren wir uns zeitnah einig, dass wir uns in unserer Datensatz-Auswahl auf dieses Thema spezialisieren wollen.

Aufgrund dieser Erfahrungen überlegten wir uns zu Beginn, inwiefern man in Abhängigkeit von den Daten Maßnahmen zur Alkoholprävention erstellen kann.

Andererseits stellte jedoch gleichermaßen die Entwicklung von Angeboten und Aktionen in Bezug auf die Alkohol-Produkte eine reizvolle Alternative dar, was auch zu unserem Studiengang „Wirtschaftsinformatik“ sehr gut passt.

Folglich trafen wir unsere Auswahl für die „*Iowa Liquor Sales*“, welche nach unserer Evaluation mit den anderen zur Verfügung stehenden Datensätzen am interessantesten auf uns wirkten und glücklicherweise aufgrund der Aktualität der Datenerhebung bislang kaum bearbeitet wurden.

## Data Background

Der ausgesuchte Datensatz stammt von <https://www.kaggle.com/>, eine Online-Plattform welche auf die Datenanalyse sowie Data Mining spezialisiert ist.

Wie der Titel schon verrät, beziehen sich die bereitgestellten Daten auf den Verkauf von hochprozentigem Alkohol im US-Bundesstaat Iowa, welche im Zeitraum zwischen dem 1. Januar und 31. Dezember 2022 erhoben und auf der landespolitischen Datenwebsite <https://data.iowa.gov/> veröffentlicht wurden.

Spezifiziert zeigt der Datensatz, zu welchem Zeitpunkt, an welchem Ort, zu welchem Preis, in welcher Menge ein bestimmtes hochprozentiges Alkoholprodukt verkauft wurde. Der Fokus liegt hierbei vor allem auf Händlern, die die sogenannte „*E*“ *liquor license* besitzen und somit jene Produkte legal verkaufen dürfen.

## Data Understanding

Aufgrund der zu Beginn präsentierten Fragestellung, ob man zu gewissen Feierlichkeiten wie Thanksgiving oder Weihnachten in Iowa bestimmte Aktionen oder Angebote in den jeweiligen Handelsläden anbietet, um so die Verkaufszahlen von Spirituosen nochmals anzukurbeln, mussten wir zuerst einmal beurteilen welche Daten überhaupt für die Beantwortung notwendig sind. Gleiches gilt natürlich auch für die Präventionsmaßnahmen.

Auch musste zunächst geprüft werden, ob die vorhandenen Daten aus den „*Iowa Liquor Sales*“ überhaupt ausreichend für die endgültige Ausarbeitung sind oder ob noch weitere Datensammlungen herangezogen werden müssen. Glücklicherweise waren aber alle notwendigen Daten vorhanden.

Allerdings wollten wir noch einige Analysen sowie Statistiken zum Alkoholkonsum pro Kopf erstellen, weshalb wir einen zweiten Datensatz hinzugezogen haben, der uns diese notwendigen Informationen über die Einwohnerzahlen liefert.

Des Weiteren wurde auch zeitnah ersichtlich, dass die Qualität der vorliegenden Daten ziemlich gut ist, da die insgesamt 2.563.482 verschiedenen Einträge, welche alle einen Kauf von hochprozentigem Alkohol widerspiegeln, mit Datum, Preis, Ort und Alkoholart vollständig versehen waren.

## Data Cleaning

Da unser Datensatz recht umfangreich war, musste zuerst die Entscheidung getroffen werden, welche Daten im Endeffekt und effektiv für die aufgestellte Fragestellung benötigt werden.

Nach dem Laden der beiden Datensätze als „*sales\_data*“ und „*population\_data*“ haben wir zu Beginn die Spalten entfernt, die für unsere Fragestellung nicht von Nöten waren.

Im Anschluss daran nutzten wir einen „*Inner Join*“ über den „Zip Code“, um die beiden Datensätze zusammenführen zu können.

Somit haben wir unseren Ausgangsdatsatz („*cleaned\_data*“) für den weiteren Verlauf erhalten, mit dem im Folgenden gearbeitet werden soll.

Letztlich sollte dieser noch auf „*nulls*“ überprüft werden, woran wir danach erkennen konnten, dass nun alle Einträge vollständig waren.

```
import pandas as pd
import plotly.express as px
import plotly.graph_objects as go
from functools import partial
import numpy as np
```

```
# Load the dataset
sales_data = pd.read_csv("sales.zip", compression='zip')
population_data = pd.read_csv("by_zip.csv", compression="infer")
print('loaded data')

# merge and drop
sales_data = sales_data.drop(['Invoice/Item Number', 'County Number', 'Iowa ZIP Code Tabulation Areas',
                             'store_number', 'Store Name', 'Address', 'Vendor Number', 'Vendor Name', 'Item Number',
                             'Iowa Watershed Sub-Basins (HUC 08)', 'Iowa Watersheds (HUC 10)', 'Pack', 'Bottle Volume (ml)',
                             'County Boundaries of Iowa', 'Store Location', 'US Counties'], axis=1)

cleaned_data = sales_data.merge(population_data, how="inner", left_on="Zip Code", right_on="zip").reset_index()

# looking for nulls
for column in ['index', 'Date', 'City', 'Zip Code', 'County', 'Category',
               'category_name', 'Item Description', 'State Bottle Cost',
               'State Bottle Retail', 'Bottles Sold', 'sale_dollars',
               'Volume Sold (Liters)', 'Volume Sold (Gallons)', 'zip', 'population']:
    print(f'{column} nulls: {cleaned_data[column].isna().sum()}')
```

## Data Handling

### Korrelationen

Zuerst starteten wir den Versuch zu analysieren, welche unserer Daten in Korrelation zueinanderstehen und erstellten in diesem Zuge eine Heatmap, um zu überprüfen ob man mit Hilfe dieser Visualisierung bereits erste Schlüsse für die weitere Ausarbeitung ziehen kann.

Leider kann man darin aber nur sehen, dass es in einzelnen Feldern starke gegenseitige Abhängigkeiten gibt, was allerdings damit begründet werden kann, dass sich diese alle auf dasselbe, nämlich die verkauften Mengen an Alkohol, beziehen.

Der einzige hierbei erkennbare Unterschied ist lediglich die Form der Angabe in Litern, Gallonen oder die Stückzahl an verkauften Flaschen.

Weiter als Ausreißer zu vermerken sind logischerweise noch die Spalten, an denen der „Inner Join“ vorgenommen wird, um die Datensätze zusammenzuführen („Zip“ und „Zip Code“).

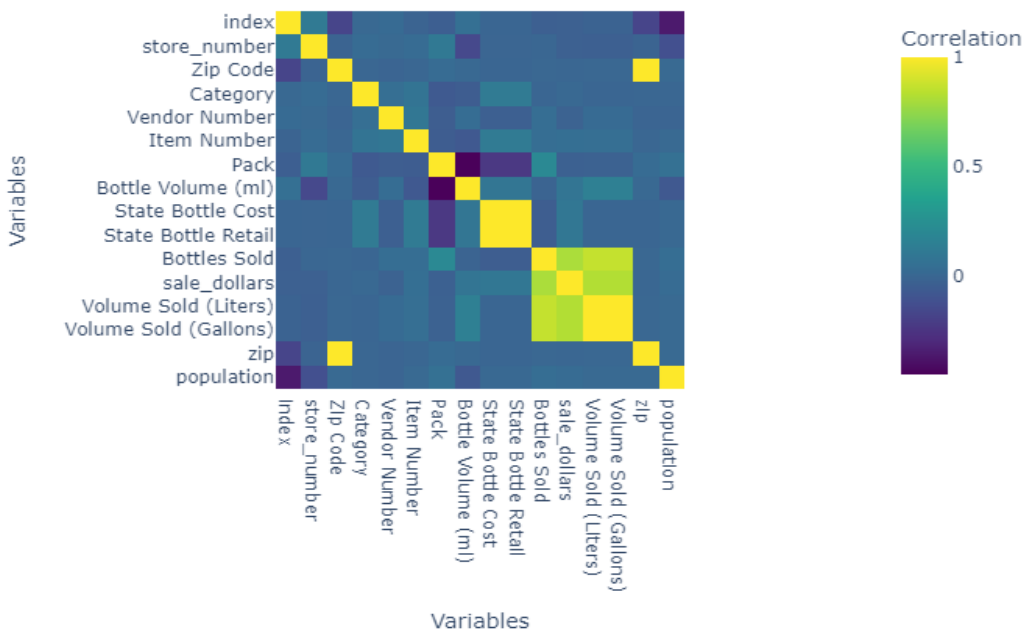
```
# Calculate correlations
numeric_columns = cleaned_data.select_dtypes(include=[np.number]).columns
correlation_matrix = cleaned_data[numeric_columns].corr()

# Create a heatmap using Plotly Express
fig = px.imshow(correlation_matrix,
                labels=dict(color="Correlation"),
                x=correlation_matrix.index,
                y=correlation_matrix.columns,
                color_continuous_scale="Viridis")

fig.update_layout(title="Correlation Matrix",
                  xaxis_title="Variables",
                  yaxis_title="Variables")

# create png
fig.write_image("heatmapAll.png")
```

Correlation Matrix



## Überlegungen zur Ausarbeitung

Zu Beginn präsentierten wir die zwei Fragestellungen, ob man mit Hilfe des ausgewählten Datensatzes Maßnahmen für die Alkoholprävention und auch aus betriebswirtschaftlicher Sicht Aktionen, beziehungsweise Angebote für Alkohol-Produkte, erstellen kann.

An diesem Punkt möchten wir uns dazu entscheiden, dass wir uns auf die betriebswirtschaftliche Alternative, die Angebote und Aktionen für Alkohol, fokussieren.

Dies ist damit begründet, weil wir zu diesem Thema mit unseren vorhandenen Daten mehrere Ergebnisse erreichen können.

Die Geodaten des Datensatzes sind teilweise unvollständig oder zu ungenau, um Alkoholpräventions-Konzepte speziell für ausgesuchte Landkreise oder Gemeinden zu erstellen, weshalb wir diese Entscheidung treffen.

Somit möchten wir uns im weiteren Verlauf der Analyse genauer die betriebswirtschaftlichen und marketingspezifischen Möglichkeiten in Bezug auf diese Thematik anschauen.

## ***Grafik 1: Übersicht über Präferenzen zu bestimmten Spirituosenarten in ganz Iowa***

Zuerst soll der Fokus darauf gerichtet werden, wie die Präferenzen der Bevölkerung zu unterschiedlichen Kategorien von Alkohol gesamtheitlich in Iowa sind.

Hierfür gruppieren wir unseren bereinigten Datensatz nach Kategorienamen und die Menge an Alkohol summiert in Litern, um feststellen zu können welche Spirituosenarten besonders beliebt sind.

Zur übersichtlicheren Darstellung für die Nutzung einer Tabelle entschieden wir uns dafür hier exemplarisch nur einen Ausschnitt von den am zehn meistverkauften Alkoholsorten einzufügen.

```
#TendencyToSpirits
print('TendencyToSpirits')
tendency_spirits = cleaned_data.groupby('category_name')['Volume Sold (Liters)'].sum().reset_index()
sorted_tendency_spirits = tendency_spirits.sort_values(by='Volume Sold (Liters)', ascending=False)

print(sorted_tendency_spirits)

fig = go.Figure(data=[go.Table(
    header=dict(values=['Category Name', 'Volume Sold (Liters)'],
                  fill_color='paleturquoise',
                  align='left'),
    cells=dict(values=[sorted_tendency_spirits['category_name'], sorted_tendency_spirits['Volume Sold (Liters)']],
                fill_color='lavender',
                align='left'))
])
fig.show()
fig.write_image("TendencyToSpiritsTabelle.png")
```

| Category Name             | Volume Sold (Liters) |
|---------------------------|----------------------|
| AMERICAN VODKAS           | 5578014.42           |
| CANADIAN WHISKIES         | 2850107.13           |
| SPICED RUM                | 1350095.25           |
| STRAIGHT BOURBON WHISKIES | 1309531.92           |
| WHISKEY LIQUEUR           | 1121244.57           |
| COCKTAILS/RTD             | 913783.54            |
| BLENDED WHISKIES          | 735642.68            |
| AMERICAN FLAVORED VODKA   | 717525.56            |
| IMPORTED VODKAS           | 699682.37            |
| 100% AGAVE TEQUILA        | 660901.85            |



Im nächsten Schritt erstellen wir nochmals eine genauere Darstellung nach Präferenzen der Alkoholart pro „Zip Code“ und sortierten diese entsprechend dem Konsum absteigend, um so eine Top zehn der „Zip Codes“ mit dem höchsten Konsum zu finden.

```
#Group and sort by Sales per Zip and Liters
grouped_data = cleaned_data.groupby(['Zip Code', 'category_name'])['Volume Sold (Liters)'].sum().reset_index()
sorted_data = grouped_data.sort_values(by=['Zip Code', 'Volume Sold (Liters)'], ascending=[True, False])
print(sorted_data)
```

## Grafik 2: „Zip Codes“ mit dem höchsten Alkoholkonsum

Um präzisieren zu können, in welchen Verkaufsmärkten unter bestimmten „Zip Codes“, das Präsentieren von entsprechenden Angebotsaktionen am effizientesten wäre, erstellten wir zwei Schaubilder.

Ersteres soll zeigen, in welchen „Zip Codes“, also Gemeinden oder Ortschaften, der generelle Absatz von hochprozentigem Alkohol in Litern am höchsten ist und letzteres wo der Verkauf/Konsum pro Kopf besonders erhöht ist.

Hierfür nahmen wir aus unserem Ausgangsdatensatz die benötigten Spalten und führten einen „*type cast*“ durch, um die Einträge auch passend verarbeiten zu können.

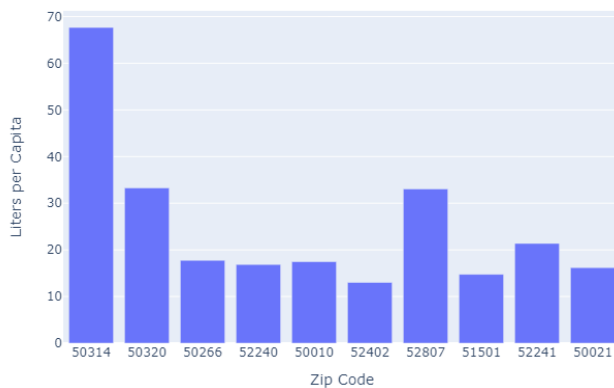
```
# LiterPerCounty.py
print('LiterPerCounty')
#LiterPerCounty
liter_per_county = cleaned_data.groupby('Zip Code')['Volume Sold (Liters)'].sum()
#Merge with another column from cleaned_data
liter_per_county = pd.merge(liter_per_county, cleaned_data[['Zip Code', 'population']].drop_duplicates(), on='Zip Code', how='left')
#TypeCasting
liter_per_county = liter_per_county.reset_index()
liter_per_county['Zip Code'] = liter_per_county['Zip Code'].astype(int)
liter_per_county['Zip Code'] = liter_per_county['Zip Code'].astype(str)
liter_per_county['Volume Sold (Liters)'] = liter_per_county['Volume Sold (Liters)'].astype(float)
liter_per_county['population'] = liter_per_county['population'].astype(float)
print(liter_per_county)
#Rechnung pro Kopf
liter_per_county['Liters per Capita'] = liter_per_county['Volume Sold (Liters)'] / liter_per_county['population']
#Sort by liters Top 10
sorted_for_liters = liter_per_county.sort_values(by='Volume Sold (Liters)', ascending=False)
top_10_county_liters = sorted_for_liters.head(10)
#Sort By Liters per Head Top 10
sorted_for_liters_per_capita = liter_per_county.sort_values(by='Liters per Capita', ascending=False)
top_10_zip_liters_per_head = sorted_for_liters_per_capita.head(10)
```

Für die Darstellungsform nutzen wir einfache Säulendiagramm.

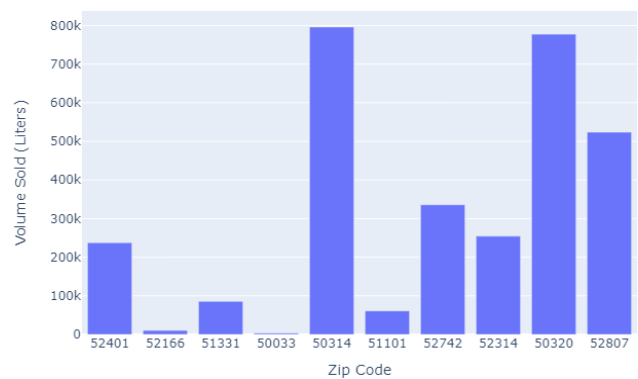
Diese besitzen den Vorteil, dass man mit ihnen sehr gut visualisieren kann, wie stark der generelle Absatz aber auch der Verkauf pro Kopf in den unterschiedlichen Gemeinden variieren kann.

```
#Graph Top10 per head
fig = px.bar(top_10_zip_liters_per_head, x='Zip Code', y='Volume Sold (Liters)', title='Volume Sold per Zip Code Top 10')
fig.show()
fig.write_image("BarGraphTop10ZipCodeLitersHead.png")
#Graph Top10 perZipCode
fig = px.bar(top_10_county_liters, x='Zip Code', y='Liters per Capita', title='Volume Sold per Head per Zip Code Top 10')
fig.show()
fig.write_image("BarGraphTop10ZipCodeLiters.png")
```

Volume Sold per Head per Zip Code Top 10



Volume Sold per Zip Code Top 10



### Grafik 3: Verkauf von Spirituosen pro Flasche über das ganze Jahr

Um genauer erkennen zu können, wie man die Daten konkret nutzen kann, um gegebenenfalls die Verkaufszahlen der Handelsläden, welche Spirituosen verkaufen, zu steigern, sollte beleuchtet werden in welcher Art und Weise sich der Verkauf von hochprozentigem Alkohol über das Jahr 2022 verhielt.

Ferner untersuchten wir, ob es auch Auffälligkeiten an bestimmten Tagen oder zu bestimmten Zeiten gibt, wann zum Beispiel besonders viel oder wenig Absatz erzielt wurde.

Hierfür erstellten wir zuerst einen Liniengraphen, der die Verkäufe pro Tag anzeigen soll, was allerdings zu einer weniger optimalen Übersicht führte.

Folgendermaßen versuchten wir die gesamten Verkäufe wochenweise und pro „Zip Code“ zusammenzufassen, da der Großteil der Bevölkerung tendenziell nur einmal in der Woche einkaufen geht, beziehungsweise Produkte wie Alkohol eher gen Wochenende gekauft werden.

Um diesen Schritt zu vollziehen, nahmen wir die nötigen Spalten („Date“, „Zip Code“, „Bottles Sold“) zusammen und summierten die Einträge der Flaschen-Verkäufe (Abbildung links, Seite 8).

Nichtsdestotrotz empfanden wir diese Darstellung immer noch als zu unübersichtlich, weshalb wir die Graphen im Folgenden auf die zehn „Zip Codes“ mit den höchsten Verkaufszahlen reduzierten, was die Darstellungsqualität leider dennoch kaum verbesserte.

Deshalb erstellten wir mit „Pandas“ einen „DataFrame“ mit den Top zehn Zip Codes und ließen diesen mit dem Ausgangsdatensatz verschmelzen.

Anschließend sollten noch „Datum“ und „Zip Code“ mit den summierten Flaschen-Verkäufen pro „Zip Code“ gruppiert werden, welche im Anschluss mit „pd.Grouper“ zu Wochen zusammengefasst und als Liniengraph ausgegeben wurde (Abbildung rechts, Seite 8).

```
# Top10 für bessere Übersicht
zips = [50314, 50320, 51501, 52240, 52402, 50266, 50010, 50613, 52404, 50317]
columns = ['Zip']
top10_zip_code_for_bottles_sold = pd.DataFrame(zips, columns=columns)
cleaned_data10 = cleaned_data.merge(top10_zip_code_for_bottles_sold, how="inner", left_on="Zip Code", right_on="Zip").reset_index()

# Group and sort by Sales per Zip and Bottle
sales_per_day_over_the_year10 = cleaned_data10.groupby(['Date', 'Zip Code'])['Bottles Sold'].sum().reset_index()

sales_per_day_over_the_year10['Date'] = pd.to_datetime(sales_per_day_over_the_year10['Date'])
sales_per_day_over_the_year10.set_index('Date', inplace=True)

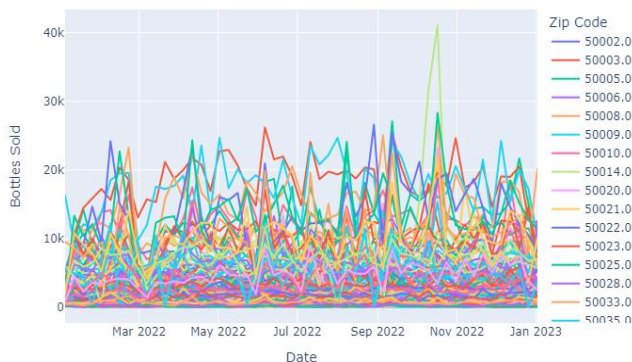
# Resample to get weekly sales per zip code
sales_per_week_over_the_year10 = sales_per_day_over_the_year10.groupby(['Zip Code', pd.Grouper(freq='W-Mon')])['Bottles Sold'].sum().reset_index()

# Create a line plot
fig = px.line(sales_per_week_over_the_year10, x='Date', y='Bottles Sold', color='Zip Code',
              title='Bottles Sold Over Time by Zip Code',
              labels={'Date': 'Date', 'Bottles Sold': 'Bottles Sold', 'Zip Code': 'Zip Code'})

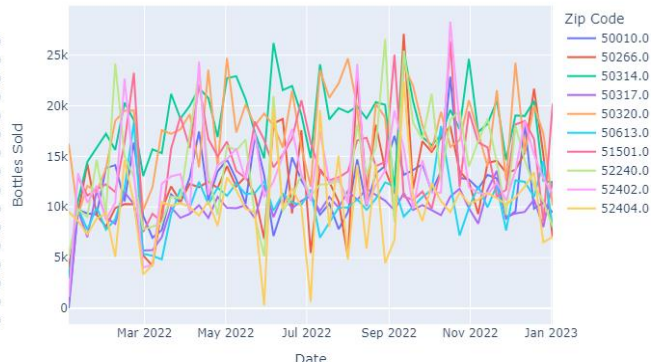
fig.show()

# create png
fig.write_image("VerkaufszahlenDateZipGraphTop10.png")
```

Bottles Sold Over Time by Zip Code



Bottles Sold Over Time by Zip Code



Für den allgemeineren Nutzen der Auswertung soll auch noch ein Liniengraph erstellt werden, um die gesamten Alkohol-Verkäufe über das gesamte Jahr 2022 in Iowa visualisieren zu können.

Hiermit hat man für die Fortsetzung der Analyse eine bessere Übersicht, mit welcher man arbeiten kann, jedoch auch gleichermaßen für die einzelnen „Zip Code“-Regionen spezialisiert passt.

```
#VerkaufszahlenDatum

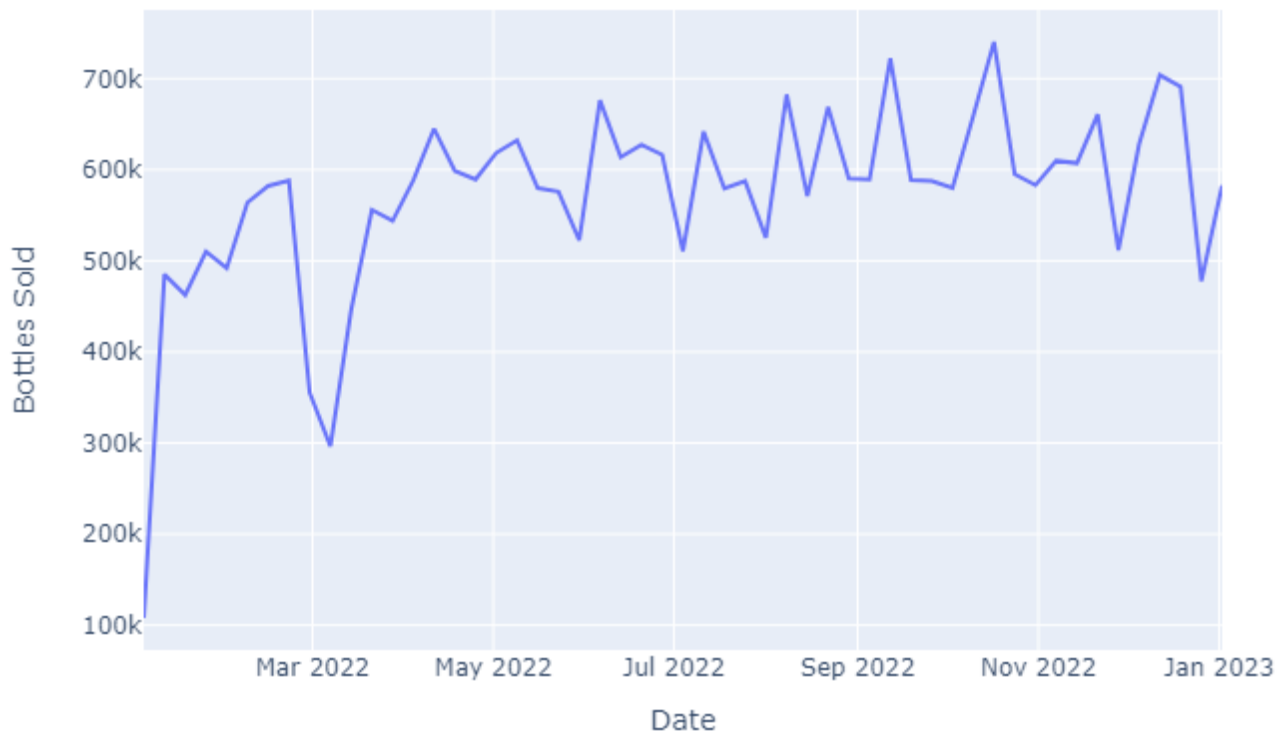
#Group and sort by Sales per Zip and Liters
sales_per_day_over_the_year = cleaned_data.groupby(['Date', 'Zip Code'])['Bottles Sold'].sum().reset_index()

sales_per_day_over_the_year['Date'] = pd.to_datetime(sales_per_day_over_the_year['Date'])
sales_per_day_over_the_year.set_index('Date', inplace=True)
sales_per_week_over_the_year = sales_per_day_over_the_year.resample('W-Mon').sum()
sales_per_week_over_the_year.reset_index(inplace=True)
sales_per_day_over_the_year.reset_index(inplace=True)

#Graph Bottles sold per week over the year
fig = px.line(sales_per_week_over_the_year, x='Date', y='Bottles Sold', title='Bottles sold over the year per week')
fig.show()
fig.write_image("SalesPerWeekOverTheYear.png")
```



Bottles sold over the year per week



Der Vorteil an diesem Liniendiagramm ist, dass man auf einen Blick sehr gut erkennen kann, wann die Menge an verkauften Flaschen besonders hoch oder gering ist. So kann man in diesem Fall sehr deutlich sehen, dass es im Zeitraum von Ende Februar bis Mitte März einen sehr deutlichen negativen Einschnitt im Graphen gibt und einen sehr hohen „Peak“ im Oktober.

Welchen Einflussfaktoren diese Ab- oder Anstiege aber unterliegen, darauf soll im nächsten Abschnitt nochmals genauer eingegangen werden.

#### ***Grafik 4: Verkäufe pro Flasche in der Woche vor Feiertagen***

Aufgrund der vorangegangenen Grafik kann man deutlich bemerken, dass es über das Jahr verteilt gewisse Schwankungen beim Absatz von Spirituosen gibt, welche näherungsweise mit einer Wellenbewegung verglichen werden können.

Allerdings gibt es, wie bereits erwähnt, besonders starke lokale Maxima und Minima.

Unsere Interpretation hierzu ist, dass dies vor allem mit den örtlichen Feiertagen zusammenhängt. So lässt sich der beachtliche negative Einschnitt im Frühjahr mit der christlichen Fastenzeit, welche vom 02.03.2022 bis zum 16.04.2022 andauerte, begründen, da zu dieser Zeit häufig auf den Konsum von Rausch- und Genussmitteln verzichtet wird.

Äquivalent hierzu können beispielsweise die Anstiege Ende Oktober und im Dezember bewertet werden. Zu diesen Zeitpunkten finden nämlich Halloween und das Weihnachtsfest statt, was in den Vereinigten Staaten beides Anlässe sind, zu denen ausgiebig gefeiert wird. Dies ist dann natürlich auch immer mit einem

höheren Absatz von alkoholischen Produkten verbunden.

Folgendermaßen erstellten wir exemplarisch hierfür zwei Tabellen, einmal für die Alkohol-Verkäufe in der Woche vor Halloween und einmal vor Weihnachten, um analysieren zu können, ob hier Differenzen im Kaufverhalten zu diesen Zeiträumen und auch im Vergleich zum Absatz der Alkoholkategorien über das gesamte Kalenderjahr 2022 (vgl. Tabelle Seite 5) erkennbar werden.

Hierfür erstellten wir zuerst die ISO-Kalenderwoche, wobei uns auffiel, dass diese ja nicht genau den Zeitraum vor den Feiertagen repräsentiert und haben deshalb selbst die Spanne an Zeit festgelegt, für welche die Daten von hier ab wieder gruppiert und sortiert wurden und schließlich in die Tabelle eingefügt werden sollten.

```
# creating 'Week' as calendar week
cleaned_data['Date'] = pd.to_datetime(cleaned_data['Date'])
cleaned_data['Week'] = cleaned_data['Date'].dt.isocalendar().week

# Filter the data for the 7 days before christmas
christmas_data = cleaned_data[((cleaned_data['Date'] < '2022-12-24') & (cleaned_data['Date'] > '2022-12-17'))]

# Group by category and sum the bottles sold
category_sales = christmas_data.groupby('category_name')['Bottles Sold'].sum().reset_index()

# Sort the data and get the top 10
top_10_categories = category_sales.nlargest(10, 'Bottles Sold')

# Create a table using plotly.graph_objects
fig = go.Figure(data=[go.Table(
    header=dict(values=['Category', 'Bottles Sold']),
    cells=dict(values=[top_10_categories['category_name'], top_10_categories['Bottles Sold'])])])

# save the table
fig.write_image("SalesBeforeChristmasTop10.png")
```

### Verkäufe vor Halloween:

| Category                       | Bottles Sold |
|--------------------------------|--------------|
| AMERICAN VODKAS                | 91470        |
| WHISKEY LIQUEUR                | 62949        |
| CANADIAN WHISKIES              | 47455        |
| STRAIGHT BOURBON WHISKIES      | 36391        |
| SPICED RUM                     | 20520        |
| 100% AGAVE TEQUILA             | 16680        |
| AMERICAN FLAVORED VODKA        | 15962        |
| MIXTO TEQUILA                  | 12014        |
| AMERICAN SCHNAPPS              | 11899        |
| TEMPORARY & SPECIALTY PACKAGES | 10912        |

### Verkäufe vor Weihnachten:

| Category                  | Bottles Sold |
|---------------------------|--------------|
| AMERICAN VODKAS           | 97393        |
| WHISKEY LIQUEUR           | 75377        |
| CANADIAN WHISKIES         | 63745        |
| STRAIGHT BOURBON WHISKIES | 35477        |
| SPICED RUM                | 26992        |
| AMERICAN SCHNAPPS         | 19481        |
| 100% AGAVE TEQUILA        | 17742        |
| BLENDED WHISKIES          | 13965        |
| AMERICAN BRANDIES         | 13910        |
| CREAM LIQUEURS            | 13601        |

Offensichtlich lässt sich feststellen, dass die Kategorien von Alkohol auf den ersten fünf Plätzen bei beiden Zeiträumen identisch sind und auch die gleiche Reihenfolge einnehmen.

Allerdings gibt es auf den unteren Plätzen auch ein paar Änderungen am Absatz von Kategorien, was es für die endgültige Auswertung sehr spannend macht, da genau diese Produkte prädestiniert dafür sind in solchen Zeitspannen per Aktionsangeboten noch stärker publiziert zu werden, um einen höheren Absatz

zu erzielen. Diese Produkte erfahren nämlich zu solchen Zeiten einen gewissen „Hype“. Wenn man dann diesen „Hype“ gezielt per Angebote befeuert, können Einzelhändler trotz niedrigerer Preise einen größeren Gewinn einfahren, da dann aufgrund der Nachfragenelastizität von solchen Gütern in der Relation zum Preisnachlass mehr gekauft wird.

Beispielhaft hierfür können für die Zeit vor Halloween die Kategorien „Mixto Tequila“ und „American Flavored Vodka“ und für die Zeit vor Weihnachten die Kategorien „Cream Liquors“ und „Blended Whiskies“ genannt werden.

## ***Ergebnisinterpretation / Erhaltenes Wissen anwenden***

### ***1. Alkoholsorten zu bestimmten Anlässen anbieten***

Abschließend kann festgehalten werden, dass mit dem bearbeiteten Datensatz und die damit erstellten Grafiken sehr gut analysiert werden kann, wie der Alkoholabsatz im Verlauf des Kalenderjahres 2022 vonstatten ging.

Auf Grundlage der Visualisierung im Liniendiagramm, welches die verkaufte Flaschen-Anzahl für das gesamte Jahr darstellt (vgl. Grafik Seite 9), kann man zu dem Schluss kommen, dass es mehrere Jahresabschnitte gibt, wie die Zeit vor Halloween oder den Weihnachtsfeiertagen, an denen es aus betriebswirtschaftlicher und auch marketingtechnischer Sicht Sinn macht, bestimmte Alkoholarten in Angebotsaktionen zu „promoten“. Wie bereits erklärt, hängt dies mit der Preiselastizität der Nachfrage dieser Produkte zusammen, da solche Güter speziell bei niedrigeren Preisen gekauft werden und man so im Vergleich zu der Preissenkung aber mehr Umsatz erzielen könnte.

So könnten zum Beispiel die Sorten „Cream Liquors“ oder „Blended Whiskies“ in der Vorweihnachtszeit mit bestimmten Rabattaktionen ausgestattet werden, weil in Folge der Datenauswertung sichtbar wurde, dass jene Sorten vermehrt in dieser Zeit verkauft wurden (vgl. Grafik Seite 10).

Andererseits würde es wenig Sinn machen, wenn man beispielsweise „Mixto Tequila“ in der Weihnachtszeit vermehrt in Rabattaktionen verwickelt, da dies ein Schnaps zu sein scheint, den die Bevölkerung traditionellerweise eher an Festtagen wie Halloween konsumiert.

Weiter interpretieren wir aus unseren Daten, dass es auch wenig Vorteile und auch keinen höheren Nutzen bringt, bestimmte Alkoholsorten während Absatzeinbrüchen, wie der bereits erwähnten Fastenzeit, anzubieten, da die Menschen unter solchen Umständen weniger Alkohol konsumieren und dementsprechend auch weniger gekauft wird (vgl. Seite 9 und 10).

## 2. Standardisierte Alkoholsorten ortsspezifisch anbieten

Im ersten Abschnitt der Ergebnisinterpretation sollten primär jene Alkoholsorten priorisiert werden, die einen stärkeren Verkaufsanstieg zu bestimmten Anlässen verzeichnen.

Allerdings ist es auch durchaus sinnvoll mehrmals im Jahr Rabattaktionen für diejenigen Alkoholsorten anzubieten, die dauerhaft in vermehrten Stückzahlen gekauft werden.

Hierzu kann man im ersten Schritt Bezug auf die Sorten nehmen, die im Jahr 2022, nach Litern bewertet, den höchsten Absatz vorweisen konnten.

Beispielhaft sind hierfür die „American Vodkas“, „Canadian Whiskies“, „Spiced Rums“ oder auch die „Straight Bourbon Whiskies“ zu nennen (vgl. Grafik Seite 5). Natürlich gibt auch bei diesen hochprozentigen Sorten gewisse Jahresabschnitte zu welchen diese stärker konsumiert werden (vgl. Grafik Seite 9). Jedoch kann bei solch marktdominanten Alkoholsorten davon ausgegangen werden, dass sie, wenn nicht immer homogen in der Menge gesehen, über das Jahr verteilt immer verkauft werden.

Dementsprechend möchten wir in unserer Interpretation und auch der finalen Auswertung genauer beleuchten, wo man jene Alkoholsorten ortsspezifisch mit Rabattaktionen versehen kann, um ebenjenen Umsatzprozess, wie im vorherigen Abschnitt erklärt, in Gang zu setzen.

Hierfür betrachten wir zuerst einmal jene „Zip Codes“, also Ortschaften, an denen besonders viel Alkoholabsatz erzielt wird (vgl. Seite 6 und 7).

Beim Betrachten der beiden Grafiken auf Seite 7 fällt auf, dass es „Zip Codes“ gibt, in welchen ein viel höherer Alkoholabsatz stattfindet als in anderen.

Dies hängt allerdings nicht damit zusammen, dass die Bevölkerung dort generell unter Äthylismus leidet, sondern dass dort schlichtweg eine größere Anzahl an Menschen lebt und dementsprechend auch die Verkaufszahlen höher sind. Des Weiteren gibt es in manchen „Zip Codes“ auch keine Einzelhandelsläden mit einer „*E*“ *liquor license*, weshalb jene Bevölkerung noch zu den anderen „Zip Codes“ in dieser Hinsicht addiert werden muss.

Jedoch würde es genau dort den größtmöglichen Nutzen erbringen, Angebotsaktionen für die oben genannten Alkoholsorten mit dem höchsten Absatz zu platzieren, da hier die größte mögliche Kundenbreite vorhanden ist.

Somit würden wir sagen, dass dies die bestmöglichen Nutzweisen sind, um das erhaltene Wissen aus dem Datensatz „*Iowa Liquor Sales*“ konkret anzuwenden: bestimmte Alkoholsorten zu bestimmten Jahresabschnitten gezielt anbieten und häufig verkaufte Alkoholsorten zusätzlich noch ortsspezifisch in Angebotsaktionen verwickeln. Dementsprechend könnte man die gewonnenen Erkenntnisse aus dem Jahr 2022 für das kommende Kalenderjahr 2024 anwenden, was den einzelnen Verkaufsläden im Bundesstaat Iowa einen deutlich höheren Umsatz wie auch Gewinn versprechen dürfte.