Viz5 - Dashboards

Hayden Ginman

2023-10-01

```
#Step 1 - Load requisite libraries into coding environment
options(repos = c(CRAN = "https://cloud.r-project.org/"))
#install.packages("scales")
#install.packages("lubridate")
library("lubridate")
```

```
##
## Attaching package: 'lubridate'
```

```
## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

```
library("scales")
library("dplyr")
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library("ggplot2")
library("reshape2")
library("readxl")
library("viridis")
```

```
## Loading required package: viridisLite
```

```
##
## Attaching package: 'viridis'
```

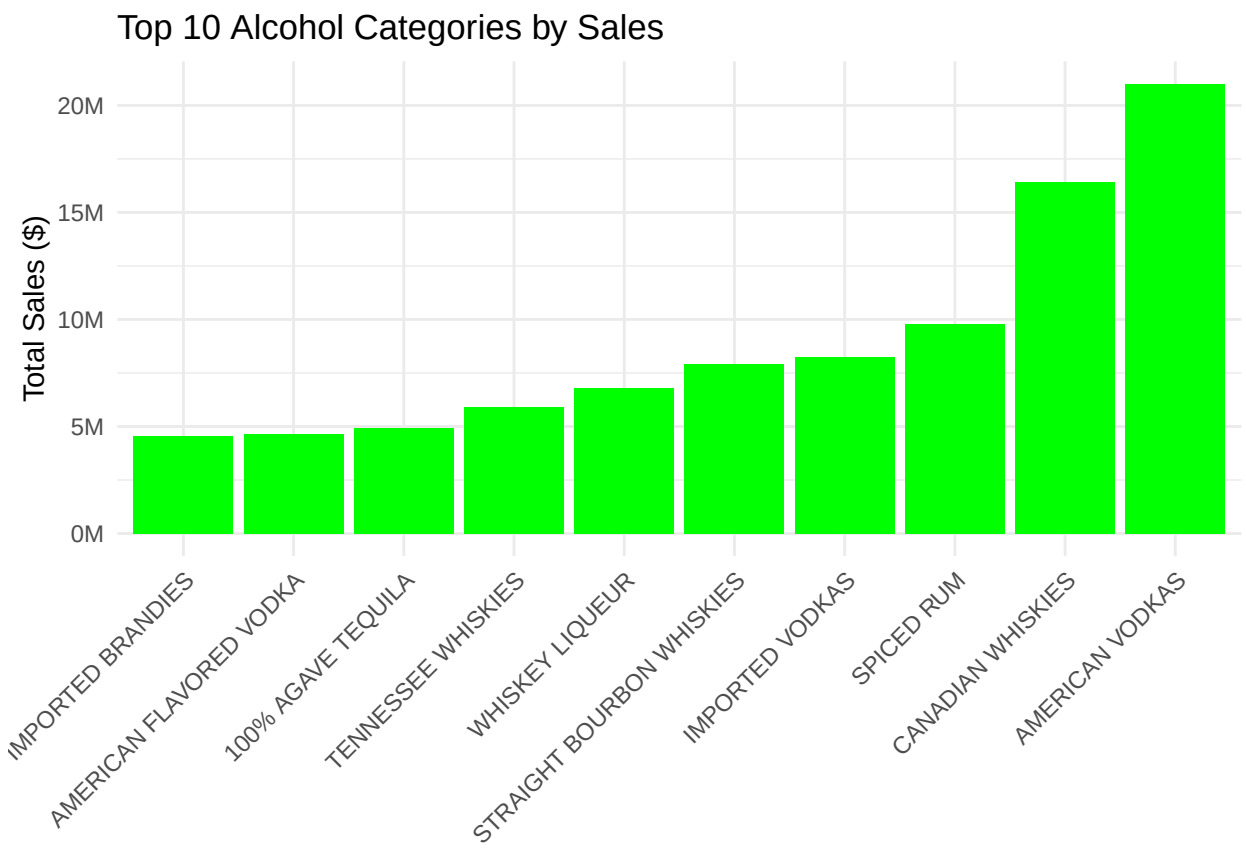
```
## The following object is masked from 'package:scales':
##
##     viridis_pal
```

```
#Step 2 - Read clean data into coding environment

iowa_sales <- read_xlsx("Smaller_Iowa_Liquor_Sales.xlsx", sheet = 1, .name_repair = "universal")

#Step 3 - Summarize the top categories and group them by category name. Summarize total sales. Do the s
```

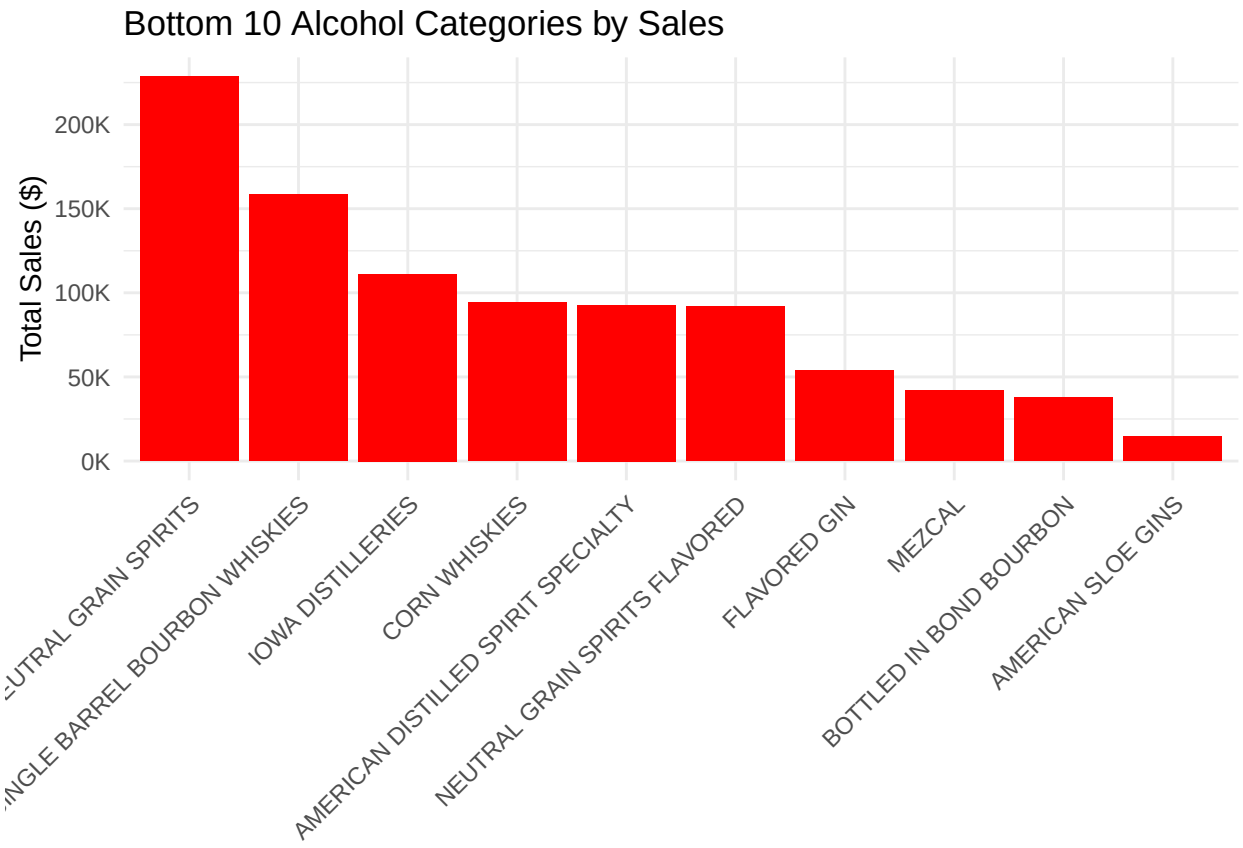
```r
top_categories <- iowa_sales %>%
   group_by(Category.Name) %>%
   summarize(TotalSales = sum(State.Bottle.Retail * Bottles.Sold))
top_categories <- top_categories %>%
   arrange(-TotalSales) %>%
   head(10)

worst_categories <- iowa_sales %>%
   group_by(Category.Name) %>%
   summarize(TotalSales = sum(State.Bottle.Retail * Bottles.Sold)) %>%
   arrange(TotalSales) %>%
   head(10)

top_categories
```

```
## # A tibble: 10 x 2
##    Category.Name             TotalSales
##    <chr>                          <dbl>
##  1 AMERICAN VODKAS            21007688.
##  2 CANADIAN WHISKIES          16438968.
##  3 SPICED RUM                  9780703.
##  4 IMPORTED VODKAS             8235447.
##  5 STRAIGHT BOURBON WHISKIES   7942365.
##  6 WHISKEY LIQUEUR             6817983.
##  7 TENNESSEE WHISKIES          5895243.
##  8 100% AGAVE TEQUILA          4924266.
##  9 AMERICAN FLAVORED VODKA     4627030.
## 10 IMPORTED BRANDIES           4564791.
```

```r
#Step 4 - Visualize the Top 10 Total sales by category

ggplot(top_categories, aes(x = reorder(Category.Name, TotalSales), y = TotalSales)) +
  geom_bar(stat = "identity", fill = "Green") +
  labs(title = "Top 10 Alcohol Categories by Sales",
       x = NULL,
       y = "Total Sales ($)") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  scale_y_continuous(labels = scales::label_number(scale = 1e-6, suffix = "M"))
```



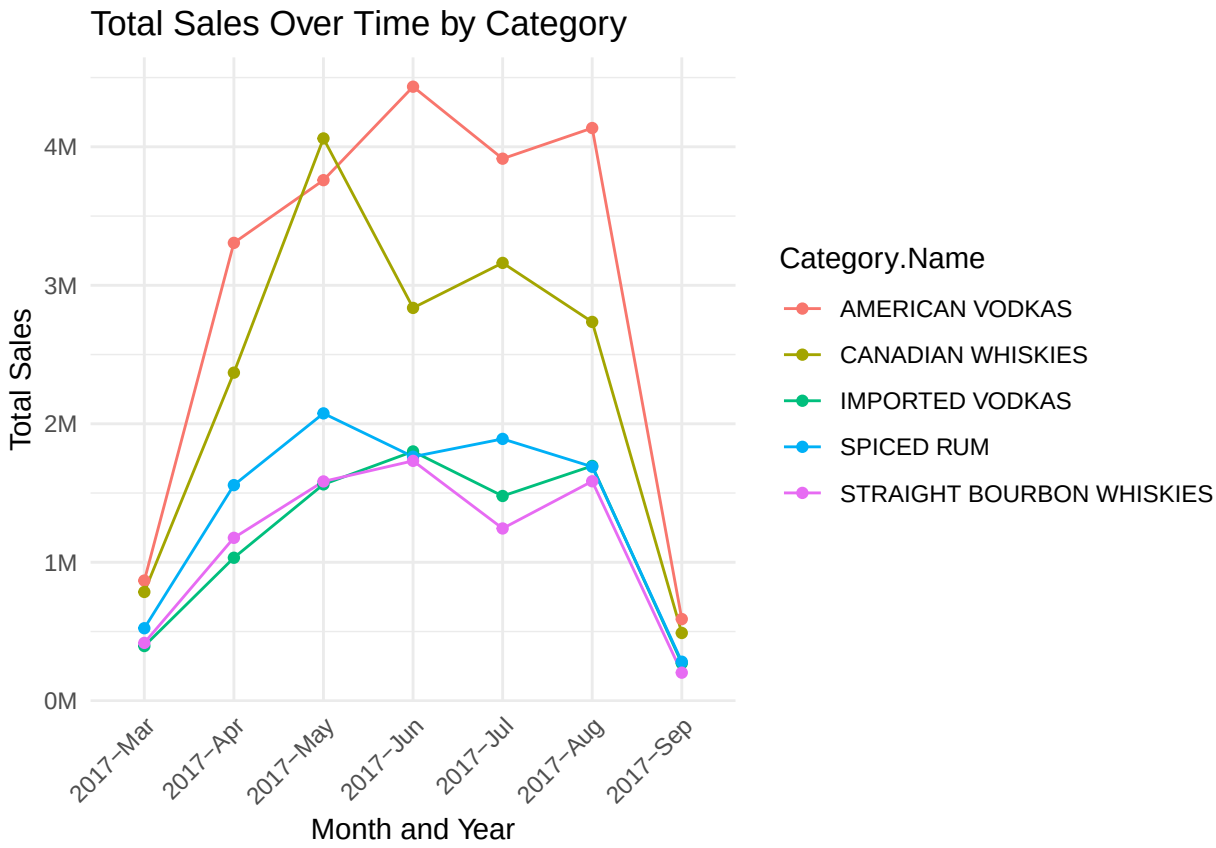

```
# Visualize the bottom 10 selling alcohol categories

ggplot(worst_categories, aes(x = reorder(Category.Name, -TotalSales), y = TotalSales)) +
  geom_bar(stat = "identity", fill = "Red") +
  labs(title = "Bottom 10 Alcohol Categories by Sales",
       x = NULL,
       y = "Total Sales ($)") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  scale_y_continuous(labels = scales::label_number(scale = 1e-3, suffix = "K"))
```




```
#Step 5 - Make line plot for sales over time for top selling categories

filtered_data <- iowa_sales %>%
  filter(Category.Name %in% c("AMERICAN VODKAS", "CANADIAN WHISKIES", "SPICED RUM", "IMPORTED VODKAS",

filtered_data$Total.Sales <- filtered_data$State.Bottle.Retail * filtered_data$Bottles.Sold

summed_sales <- filtered_data %>%
  mutate(Year = year(Date), Month = month(Date, label = TRUE)) %>%
  group_by(Year, Month, Category.Name) %>%
  summarise(Total.Sales = sum(Total.Sales))
```

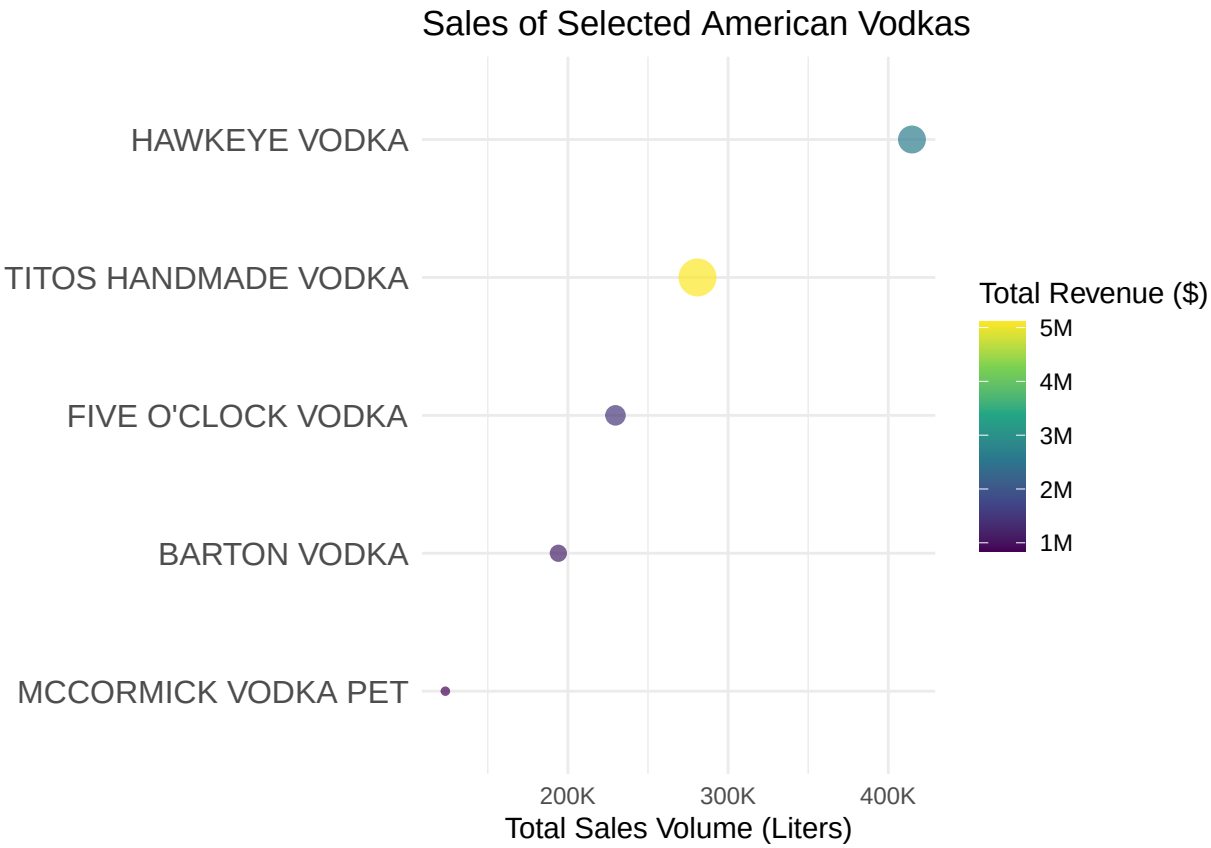
```
## `summarise()` has grouped output by 'Year', 'Month'. You can override using the
## `.groups` argument.
```

```
ggplot(summed_sales, aes(x = interaction(Year, Month, sep = "-"), y = Total.Sales, color = Category.Nam
  geom_line(aes(group = Category.Name)) +
  geom_point() +
  labs(title = "Total Sales Over Time by Category", x = "Month and Year", y = "Total Sales") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  scale_y_continuous(labels = scales::label_number(scale = 1e-6, suffix = "M"))
```




```r
# Make visualization for best selling American Vodkas

american_vodka_sales <- iowa_sales %>%
  filter(Category.Name == "AMERICAN VODKAS") %>%
  group_by(Item.Description) %>%
  summarize(
    TotalVolume = sum(Volume.Sold.Liters),
    TotalRevenue = sum(Bottles.Sold * State.Bottle.Retail)
  ) %>%
  arrange(-TotalVolume)

selected_vodkas <- c("HAWKEYE VODKA", "TITOS HANDMADE VODKA", "FIVE O'CLOCK VODKA", "BARTON VODKA", "MC

filtered_vodka_sales <- american_vodka_sales %>%
  filter(Item.Description %in% selected_vodkas)

ggplot(filtered_vodka_sales, aes(x = reorder(Item.Description, TotalVolume), y = TotalVolume)) +
  geom_point(aes(size = TotalRevenue, color = TotalRevenue), alpha = 0.7) +
  coord_flip() +
  theme_minimal() +
  labs(
    title = "Sales of Selected American Vodkas",
    x = NULL,
    y = "Total Sales Volume (Liters)") +
  theme(axis.text.y = element_text(size = 12)) +
  guides(size = "none") +
  scale_color_viridis_c(name = "Total Revenue ($)", labels = scales::label_number(scale = 1e-6, suffix =
```

```
  scale_y_continuous(labels = scales::label_number(scale = 1e-3, suffix = "K"))
```




```
# Calculate profit margin for best selling vodkas and visualize it

profit_margin_data <- iowa_sales %>%
  filter(Item.Description %in% selected_vodkas) %>%
  group_by(Item.Description) %>%
  summarize(
    TotalRevenue = sum(State.Bottle.Retail * Bottles.Sold),
    TotalCost = sum(State.Bottle.Cost * Bottles.Sold)) %>%
  mutate(ProfitMargin = (TotalRevenue - TotalCost) / TotalRevenue * 100)

profit_margin_data
```

```
## # A tibble: 5 x 4
##   Item.Description     TotalRevenue TotalCost ProfitMargin
##   <chr>                       <dbl>     <dbl>        <dbl>
## 1 BARTON VODKA             1163141.   775014.         33.4
## 2 FIVE O'CLOCK VODKA       1491609.   993805.         33.4
## 3 HAWKEYE VODKA            2637605.  1757743.         33.4
## 4 MCCORMICK VODKA PET       843175.   562006.         33.3
## 5 TITOS HANDMADE VODKA     5111072.  3407003.         33.3
```

```
#Make plot for alcohol that generates the least amount of revenue

iowa_sales <- iowa_sales %>%
  mutate(Revenue = State.Bottle.Retail * Bottles.Sold)
```

```r
least_revenue_alcohol <- iowa_sales %>%
  group_by(Item.Description) %>%
  summarize(TotalRevenue = sum(Revenue)) %>%
  arrange(TotalRevenue)

least_revenue_alcohol <- head(least_revenue_alcohol, 20)

ggplot(least_revenue_alcohol, aes(x = Item.Description, y = TotalRevenue)) +
  geom_point(color = "red") +
  coord_flip() +
  labs(
    title = "Alcohol Products with Least Revenue",
    x = "Item Description",
    y = "Total Revenue"
  ) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1))
```