Springboard’s Foundations of Data Science

Stephanie Vaul

Capstone Project Proposal Ideas

1. **BMW Sales Trends in the U.S.**

Review BMW Sales data to determine trends in purchasing behavior, including the following:

* Look for brand loyalty by finding repeat buyers via multiple purchase records, as well as traded-in data.
  + Determine which model series are the most popular for repeat buyers.
  + Determine which model series are the most popular for 1st time buyers.
  + Determine which regions have the most buyers – new and repeat.
* Look for trends among different model series (1-series, 3-series, 5-series, etc)
  + by regional
  + by economic status

This exploratory analysis will help BMW and its marketing partners better focus their sales marketing efforts by leveraging existing trends, as well as being able to extrapolate those trends in other areas of similar makeup.

Data Sources:

1. Proprietary sales data from most BMW centers in the U.S. going back 10yrs

PurchDate

VehYear

VehModel

TradeVIN

TradeVehYear

TradeVehModel

CustState – to link back to economic factors

CustZip – to link back to economic factors

PurchType (Lease, Cash, Finance)

CustID – to look for repeat purchases

1. Economic data (see potential data sets below)
2. **U.S. Vehicle Sales Trends across OEMs** (Original Equipment Manufacturer – i.e. Vehicle Makes)

Determine which automotive OEMs an automotive marketing company should target exclusive OEM marketing contracts in order to gain the most profitability & future growth.

Look for over trends by automotive OEM sales in the U.S. to determine which manufactures are increasing or decreasing in sales over the last decade (more or less depending on availability). While growth rate is one factor, we’ll also want to look at overall sales numbers because even if a luxury vehicle, like Aston Martin or Rolls-Royce, is experiencing increased popularity, it may not have enough overall vehicle sales in the U.S. to support the development & cost required to capture & manage that OEM’s exclusive marketing business. Also looking for trend by region (i.e. West, Central/North, South, & Eastern) & possibly by economic status within each Region. Even without exclusive contracts, it is still possible to target dealerships within the U.S. or a particular region, so we’ll want to know where to put our focus.

**Potential Data Sources:**

Quandl - Economic Data

<https://www.quandl.com/browse?idx=database-browser_economic-data_united-states_inflation-and-price-data>

Economic Data / US / **Economy and Society Data** /

Interest Rate Data / US / **Benchmark Interest Rates - &- LIBOR and Swap Rates /**

**Federal Reserve Economic Data**

<https://www.quandl.com/data/FRED?keyword>=

Growth, employment, inflation, labor, manufacturing and other US economic statistics from the research department of the Federal Reserve Bank of St. Louis.

* GNP
* <https://www.quandl.com/data/FRED/NGDPPOT-Nominal-Potential-Gross-Domestic-Product>
* Unemployment – Short-Term
* <https://www.quandl.com/data/FRED/NROUST-Natural-Rate-of-Unemployment-Short-Term>
* Unemployment – Long-Term
* <https://www.quandl.com/data/FRED/NROU-Natural-Rate-of-Unemployment-Long-Term>

Interest Rate Data / US / **Benchmark Interest Rates /**

**Federal Reserve Bank of Cleveland**

<https://www.quandl.com/data/FRBC?keyword>=

The Federal Reserve Bank of Cleveland collects data from hundreds of financial institutions, including depository institutions, bank holding companies, and other entities that is used to assess financial institution conditions and also to glean insights into how the economy and financial system are doing.

Interest Rate Data / US / **Par, Zero-Coupon and TIPS Yield Curves** /

Economic Data / US / **Banking, Finance and Monetary Data** /

**US Federal Reserve Data Releases**

<https://www.quandl.com/data/FED?keyword>=

Official US figures on money supply, interest rates, mortgages, government finances, bank assets and debt, exchange rates, industrial production.

Economic Data / US / **Economy and Society Data** /

**Thomas Pikentty**

<https://www.quandl.com/data/PIKETTY?keyword>=

Data on Income and Wealth from "Capital in the 21st Century", Harvard University Press 2014.

Economic Data / US / **Labor, Employment and Productivity Data** /

**BLS Employment & Unemployment**

<https://www.quandl.com/data/BLSE?keyword>=

US national and state-level employment and unemployment statistics, published by the Bureau of Labor Statistics.

Economic Data / US / **Inflation and Price Data** /

**BLS Inflation & Prices**

<https://www.quandl.com/data/BLSI?keyword>=

US national and state-level inflation data, published by the Bureau of Labor Statistics.

Economic Data / US / **Government, Census and Tax Data** /

**U.S. Census Bureau**

<https://www.quandl.com/data/USCENSUS?keyword>=

Data on the American people, places and economy. It provides many data on U.S. imports/exports, domestic production, and other key national indicators.

**US Treasury**

<https://www.quandl.com/data/USTREASURY?keyword>=

The U.S. Treasury ensures the nation's financial security, manages the nation's debt, collects tax revenues, and issues currency, provides data on yield rates.

**US Gov Data**

Annual Retail Trade Survey

<http://catalog.data.gov/dataset/annual-retail-trade-survey>

national estimates of total annual sales, e-commerce sales, end-of-year inventories, inventory-to-sales ratios, purchases, total operating expenses, inventories held outside the United States, gross margins, and end-of-year accounts receivable for retail businesses and annual sales and e-commerce sales for accommodation and food service firms located in the U.S.

[State Government Tax Collections](http://catalog.data.gov/dataset/state-government-tax-collections)

<http://catalog.data.gov/dataset/state-government-tax-collections>

Department of Commerce —

The State Government Tax Collections report provides a summary of taxes collected by state for up to 25 tax categories. These tables and data files present the...

[Manufacturing & Trade Inventories & Sales](http://catalog.data.gov/dataset/manufacturing-trade-inventories-sales)

<http://catalog.data.gov/dataset/manufacturing-trade-inventories-sales>

Department of Commerce —

To provide broad and timely measures of combined changes in business sales and end-of-month inventories for domestic retail trade, wholesale trade and...

[Annual Wholesale Trade](http://catalog.data.gov/dataset/annual-wholesale-trade)

<http://catalog.data.gov/dataset/annual-wholesale-trade>

Department of Commerce —

Provides estimates on annual sales, end-of-year inventories, inventory valuation, purchases, operating expenses and e-commerce data for merchant wholesalers and...

[MVA Vehicle Sales Counts by Month for CY 2002 - 2015](http://catalog.data.gov/dataset/mva-vehicle-sales-counts-by-month-for-cy-2002-2015)

<http://catalog.data.gov/dataset/mva-vehicle-sales-counts-by-month-for-cy-2002-2015>

State of Maryland —

The number of new and used vehicles and the sales dollars respectively sold by month. Data as of October 2015

[Advanced Monthly Retail Trade Survey](http://catalog.data.gov/dataset/advanced-monthly-retail-trade-survey)

<http://catalog.data.gov/dataset/advanced-monthly-retail-trade-survey>

Department of Commerce —

Provides an early indication of sales of retail and food service companies throughout the United States.

Google: “us vehicle sales by model”

<https://ycharts.com/indicators/auto_sales>

**US Vehicle Sales View and export this data going back to 1976. Source:** [Bureau of Economic Analysis](https://ycharts.com/indicators/sources/bea)

“Access decades of history for stock prices, company financial metrics, economic data and more. Start your free 7-Day Trial.”

*not sure if this shows by OEM or just monthly totals…*

GoodCarBadCar – has annual sales changes for several Make/Models over a one year period (i.e. June 2015 vs June 2016) for both YTD as well as just that month. Also have June 2014 vs June 2015, as well as June 2013 vs June 2014 – so three years’ worth of sales by most popular make/models.

<http://www.goodcarbadcar.net/2016/07/usa-vehicle-sales-by-model-june-2016-first-half.html>

<http://www.goodcarbadcar.net/2015/07/usa-june-2015-ytd-auto-sales-figures-by-model.html>

<http://www.goodcarbadcar.net/2014/07/usa-all-vehicle-nameplates-ranked-june-2014-ytd-sales-figures.html>

<http://www.goodcarbadcar.net/2013/07/usa-auto-sales-rankings-by-model-june-2013-ytd.html>

GoodCarBadCar also has some total market figures that can be queried for the U.S.

<http://www.goodcarbadcar.net/p/sales-stats.html>

Overall Market goes back to 1999 for Annual Sales & back to Jan 2010 for Monthly sales.

The same can be queried for any OEM and model series so potentially, I can grab 6 ½ years’ worth of U.S. Sales data by Month or 10+ years’ worth annually. Overall max goes back to 1999, but other brands may have less data. For example, BMW & Ford brands only go back to 2002.

All of their data appears to copy and paste well into Excel, although some do merge cells.

Ex of JUNE to JUNE Top Sales data

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Rank** | **Vehicle** | **June** | **June** | **%** | **June** | **June** | **% Change** |
| **2014** | **2013** | **Change** | **2014** | **2013** |
| **YTD** | **YTD** |  |  |  |
| #1 | Ford F-Series | 365,825 | 367,486 | -0.50% | 60,560 | 68,009 | -11.00% |
| #2 | Chevrolet Silverado | 240,679 | 242,586 | -0.80% | 43,519 | 43,259 | 0.60% |
| #3 | Toyota Camry | 222,540 | 207,626 | 7.20% | 40,664 | 35,870 | 13.40% |
| #4 | Ram P/U | 203,860 | 170,319 | 19.70% | 33,149 | 29,644 | 11.80% |
| #5 | Honda Accord | 185,278 | 186,860 | -0.80% | 32,329 | 31,677 | 2.10% |
| #6 | Nissan Altima | 176,453 | 167,787 | 5.20% | 26,111 | 26,904 | -2.90% |
| #7 | Toyota Corolla/Matrix | 174,354 | 158,972 | 9.70% | 30,945 | 26,458 | 17.00% |
| #8 | Honda Civic | 167,097 | 158,704 | 5.30% | 32,301 | 29,724 | 8.70% |
| #9 | Ford Fusion | 165,498 | 161,146 | 2.70% | 27,604 | 24,313 | 13.50% |
| #10 | Honda CR-V | 154,692 | 145,763 | 6.10% | 26,129 | 26,572 | -1.70% |
| #11 | Ford Escape | 152,890 | 156,626 | -2.40% | 25,110 | 28,694 | -12.50% |

Example of Ford Brand Query – sales by month

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Month** | **Ford** | **Ford** | **Ford** | **Ford** | **Ford** |  |  |
| **U.S. Sales 2010** | **U.S. Sales 2011** | **U.S. Sales 2012** | **U.S. Sales 2013** | **U.S. Sales 2014** | **Ford U.S. Sales 2015** | **Ford U.S. Sales 2016** |
| **January** | 99,631 | 121,175 | 131,173 | 161,672 | 147,521 | 170,822 | 165,301 |
| **February** | 123,228 | 150,284 | 171,732 | 190,427 | 176,688 | 173,509 | 208,006 |
| **March** | 159,009 | 203,794 | 214,081 | 228,818 | 234,448 | 226,091 | 243,375 |
| **April** | 146,330 | 182,048 | 173,350 | 204,369 | 203,552 | 213,518 | 219,963 |
| **May** | 175,129 | 184,130 | 208,425 | 238,714 | 244,501 | 240,912 | 224,941 |
| **June** | 155,127 | 186,054 | 199,660 | 227,448 | 214,125 | 216,355 | 230,287 |
| **July** | 153,400 | 172,501 | 166,507 | 186,161 | 203,604 | 212,478 | 206,170 |
| **August** | 143,859 | 166,794 | 188,608 | 212,212 | 213,227 | 225,244 |  |
| **September** | 146,559 | 167,842 | 167,652 | 177,999 | 172,261 | 212,589 |  |
| **October** | 143,339 | 161,408 | 162,793 | 184,136 | 179,014 | 204,620 |  |
| **November** | 133,162 | 160,136 | 171,360 | 182,978 | 178,221 | 178,971 |  |
| **December** | 173,738 | 201,044 | 205,518 | 208,608 | 209,679 | 226,746 |  |

Example of Ford Brand Query – sales by year

|  |  |
| --- | --- |
| **Year** | **Ford Brand** |
| **U.S. Sales** |
| **2002** | 2,990,472 |
| **2003** | 2,886,575 |
| **2004** | 2,766,169 |
| **2005** | 2,634,041 |
| **2006** | 2,415,059 |
| **2007** | 2,087,048 |
| **2008** | 1,680,321 |
| **2009** | 1,440,653 |
| **2010** | 1,752,511 |
| **2011** | 2,057,210 |
| **2012** | 2,160,859 |
| **2013** | 2,403,542 |
| **2014** | 2,376,841 |
| **2015** | 2,501,855 |
| **2016 YTD \*** | 1,498,043 |

\*through July 2016

**The Wall Street Journal -** This site has a lot of analysis already performed on different segments of brands and types (cars, trucks, SUV's, etc) from 2015 to 2016.

<http://online.wsj.com/mdc/public/page/2_3022-autosales.html>

**Trading Economics –** This may be a good site to explore further; however, their “download data” takes you to a subscription page with the lowest price of $49 for limited data retrieval for one week, then $199 each month. So it’s a no go.

<http://www.tradingeconomics.com/united-states/total-vehicle-sales>