

# Data preparation

ANALYZING DATA IN TABLEAU



**Lis Sulmont**

Head of Curriculum at DataCamp

# Data preparation

Ask yourself...

- Do any fields need to be refined?
- Are there calculated fields we can create to more effectively tell our data story?
- Does the data contain fields that will allow for summaries or grouping at a higher level?
- Are there sufficient categorical fields to **slice and dice** your data?



<sup>1</sup> Photo credit: Arvell Dorsey Jr. from Chicago, IL, United States

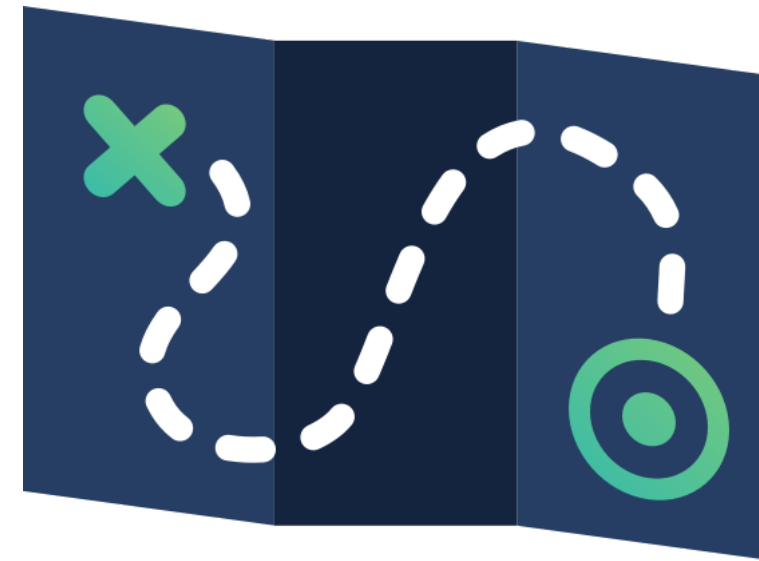
# Divvy dataset: stations table



- `id` : ID attached to each station
- `name` : station name
- `latitude` : station latitude
- `longitude` : station longitude
- `docks` : number of docks at the station

# Divvy dataset: trips table

- *Trips taken between Jan - June, 2019*
- `trip id` : ID attached to each trip
- `bikeid` : ID attached to each bike
- `tripduration` : time of trip in seconds
- `starttime` : day and time trip started (CST)
- `endtime` : day and time trip ended (CST)
- `from station id` : station ID of trip start
- `from_station_name` : station name of start
- `to station id` : station ID of trip end



- `to station name` : station name of end
- `usertype` : *customer* or *subscriber*
- `birthyear` : birth year of rider
- `gender` : gender of rider

# Dimension and measure recap

## Dimensions:

- Categorical or qualitative data

## Measures:

- Numerical data that can be aggregated

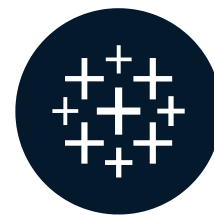
**We want to move fields strategically between these two types:**

- Move numeric fields that shouldn't be aggregated to the Dimensions section

**Let's practice!**  
ANALYZING DATA IN TABLEAU

# Preparing the data

ANALYZING DATA IN TABLEAU



**Lis Sulmont**

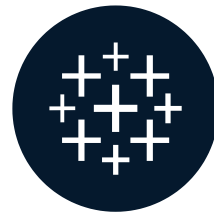
Head of Curriculum at DataCamp



**Let's practice!**  
ANALYZING DATA IN TABLEAU

# Calculated Fields to extend data

ANALYZING DATA IN TABLEAU



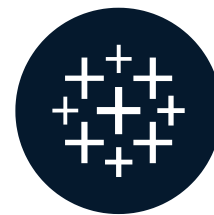
**Lis Sulmont**

Head of Curriculum at DataCamp

**Let's practice!**  
ANALYZING DATA IN TABLEAU

# Visualizations for exploratory analysis of trends

ANALYZING DATA IN TABLEAU



**Lis Sulmont**

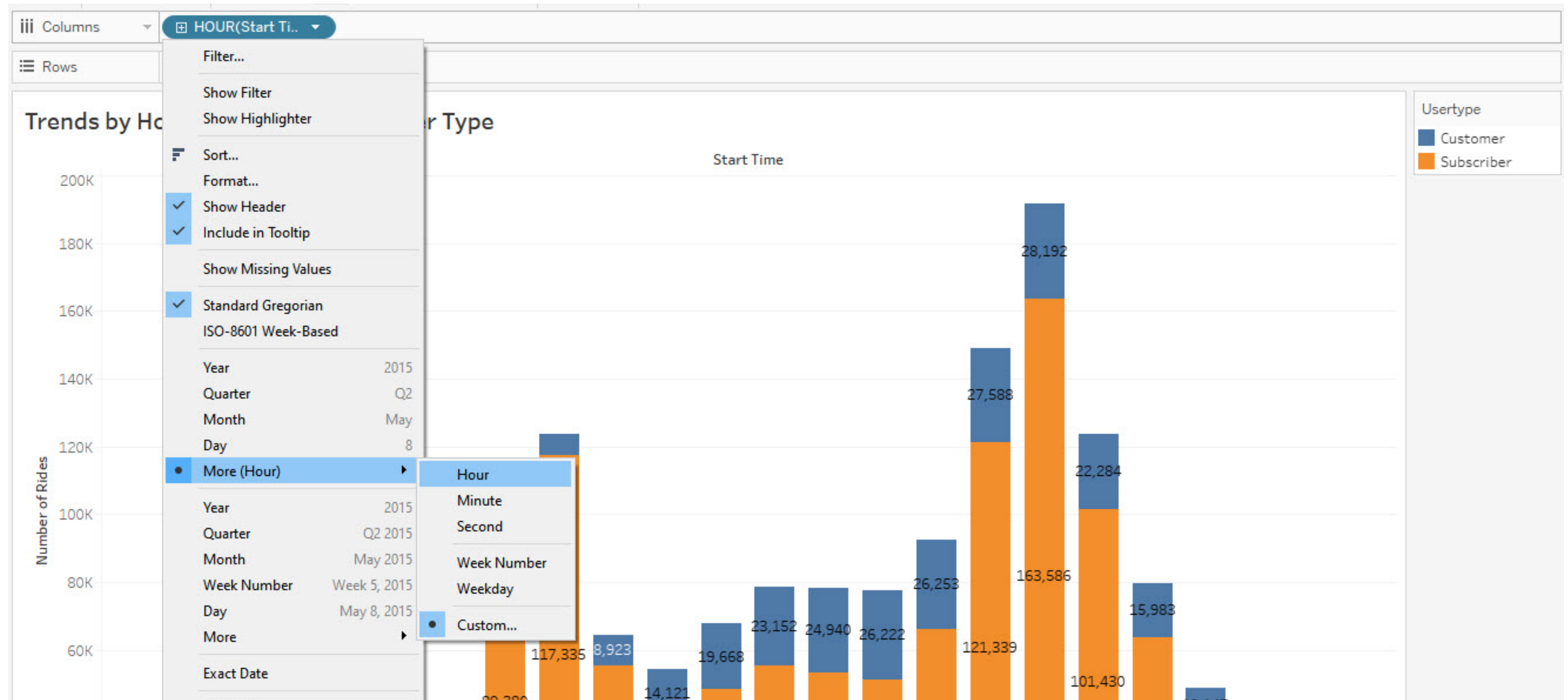
Head of Curriculum at DataCamp

# Looking for trends



- Marketing opportunities
- Scheduling maintenance
- Managing size and scheduling of staff
- Increasing or decreasing product stock or availability
- Hourly, daily monthly, annual

# Configuring data on Tableau



# Discrete or continuous time analysis?

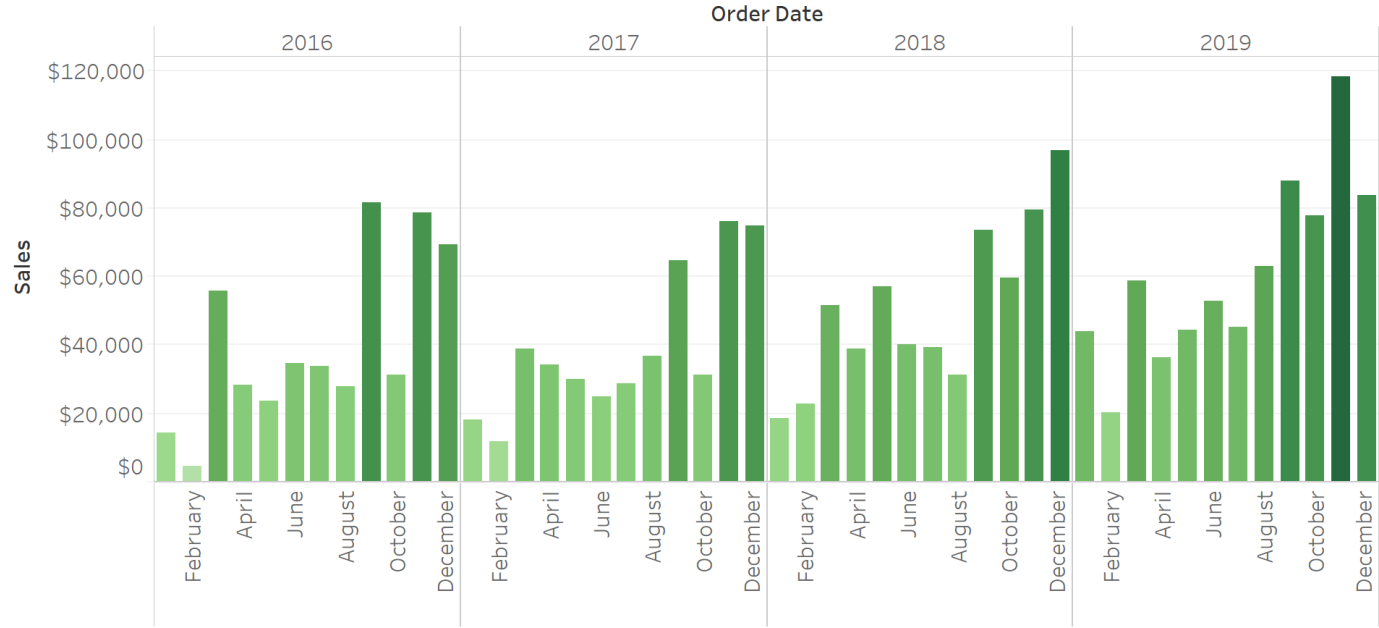
## Discrete (bins):

- Trends by hour, day of the week, month, etc

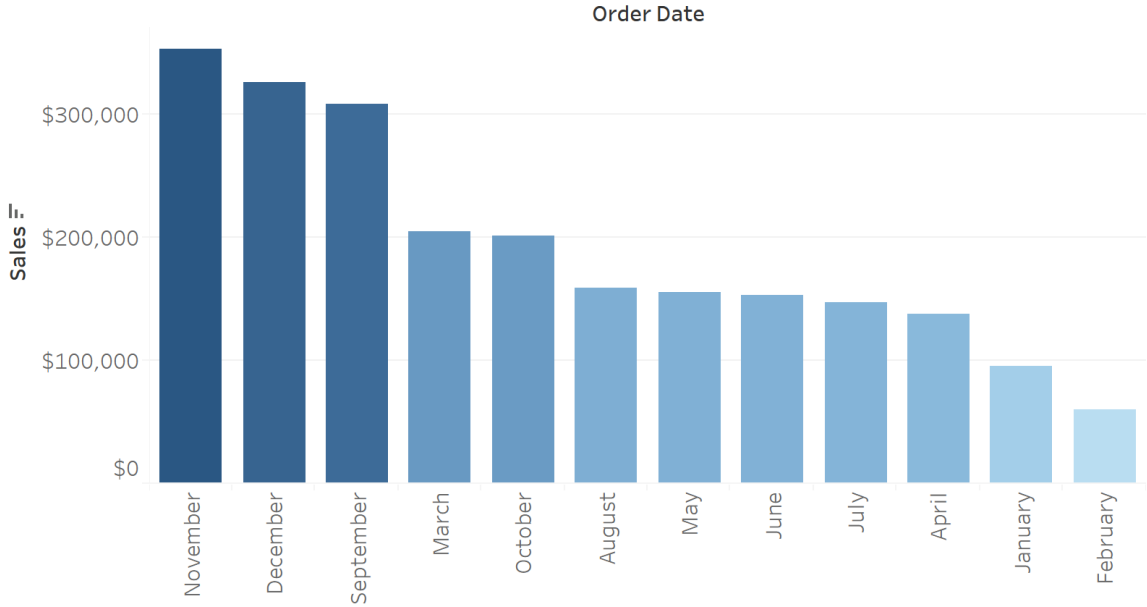
## Continuous (time series):

- Presenting data over time in the sequence it historically occurred

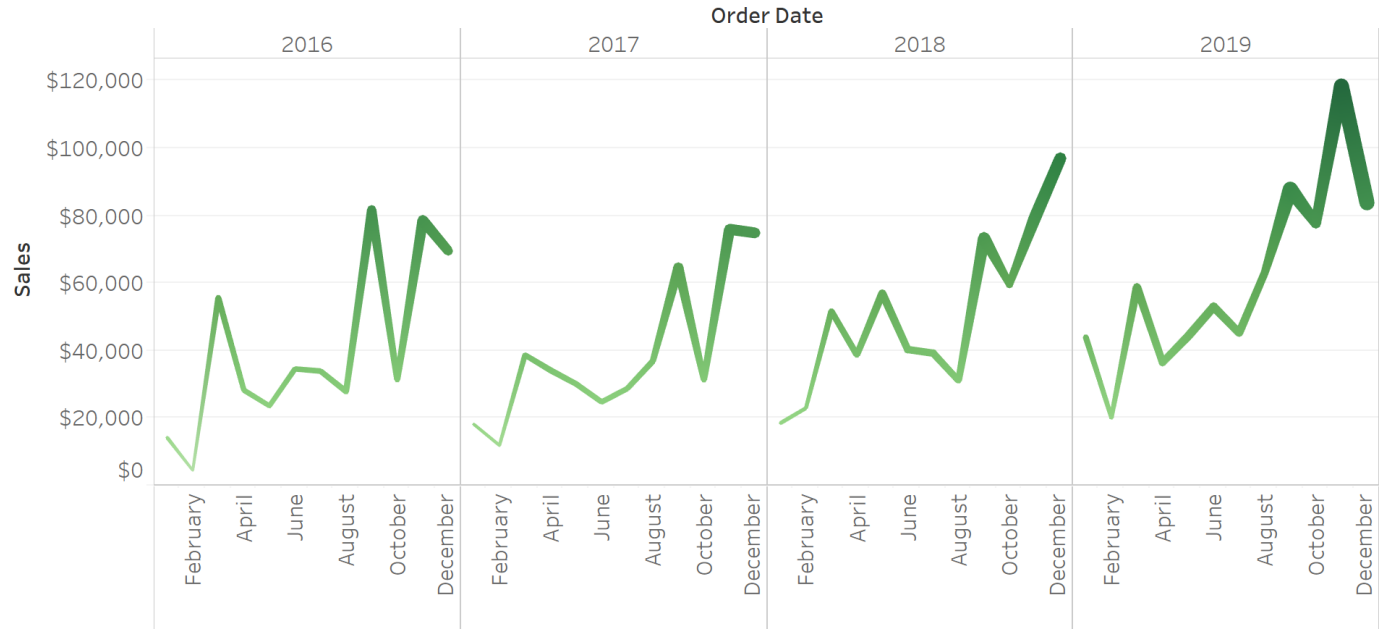
# Trends Over Time (Barchart)



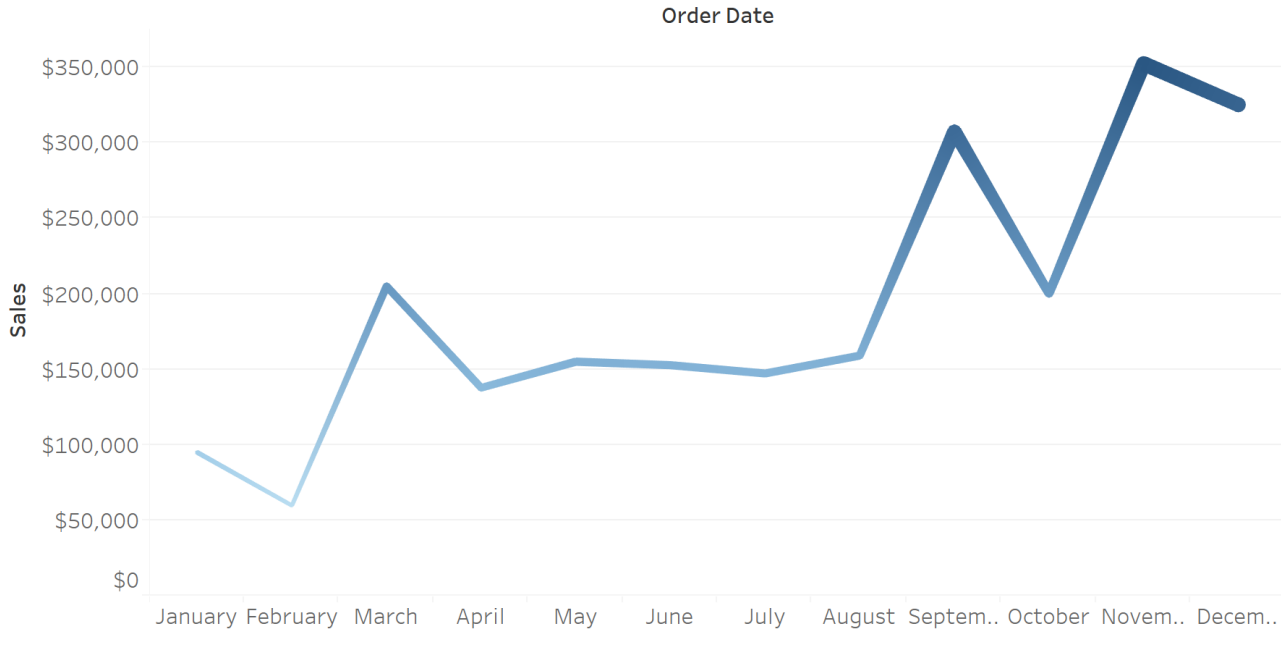
# Monthly Pattern of Sales Trends



# Trends Over Time (Line Chart)

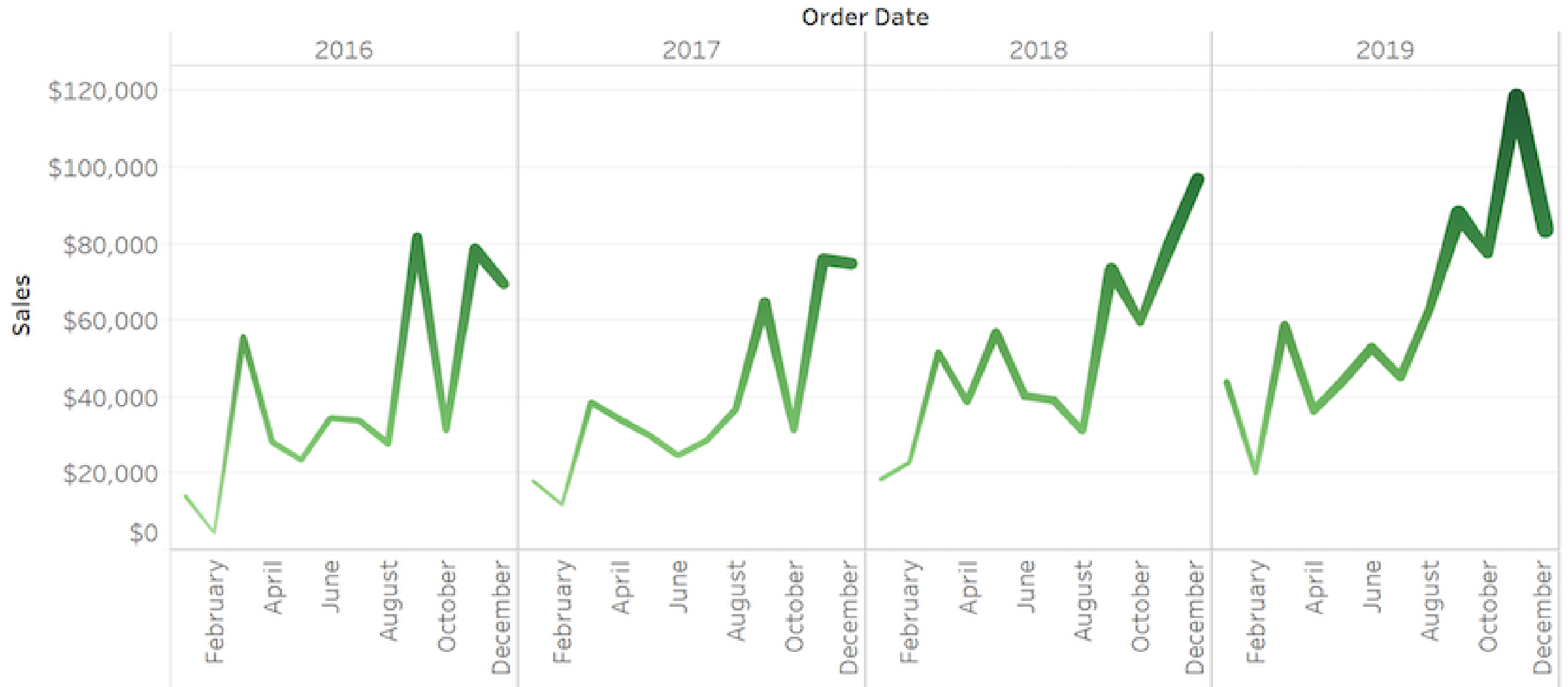


# Monthly Trendline of Sales Trends

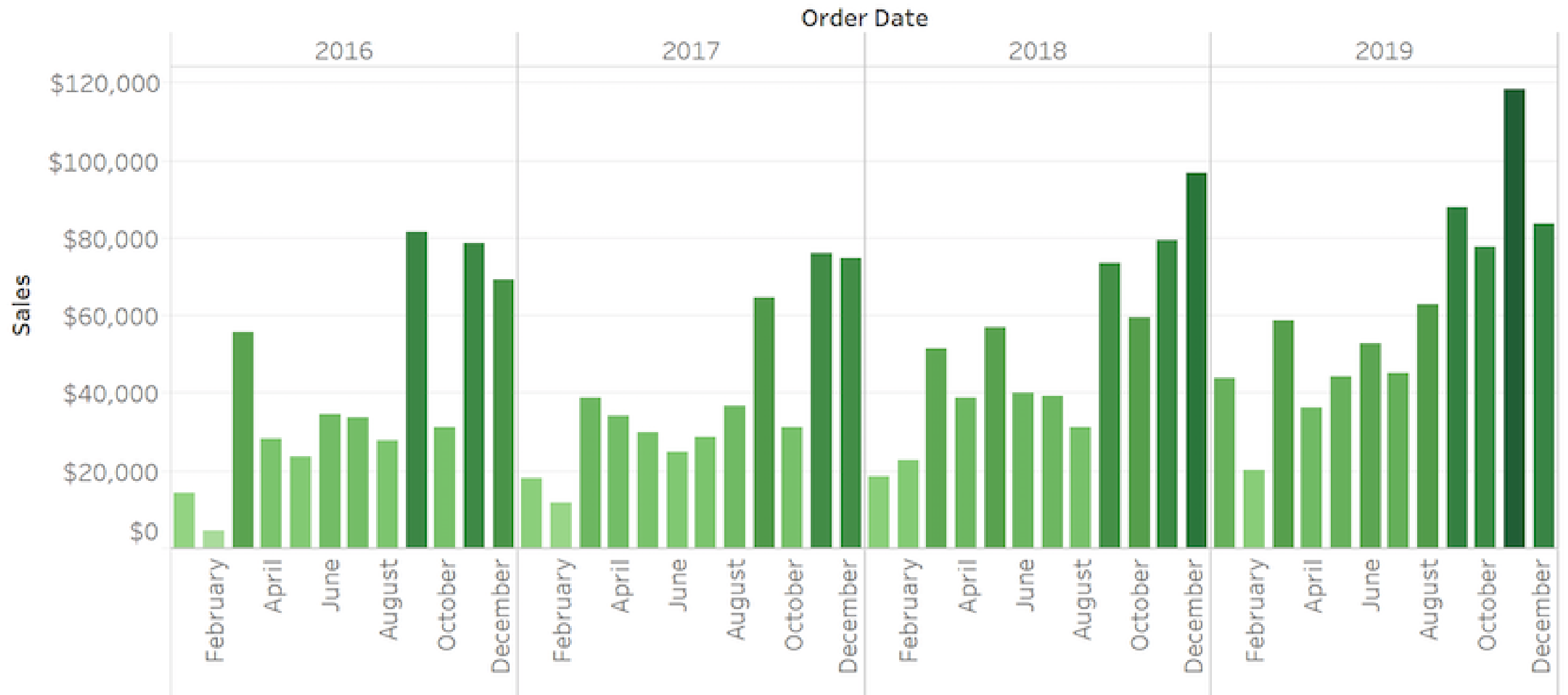




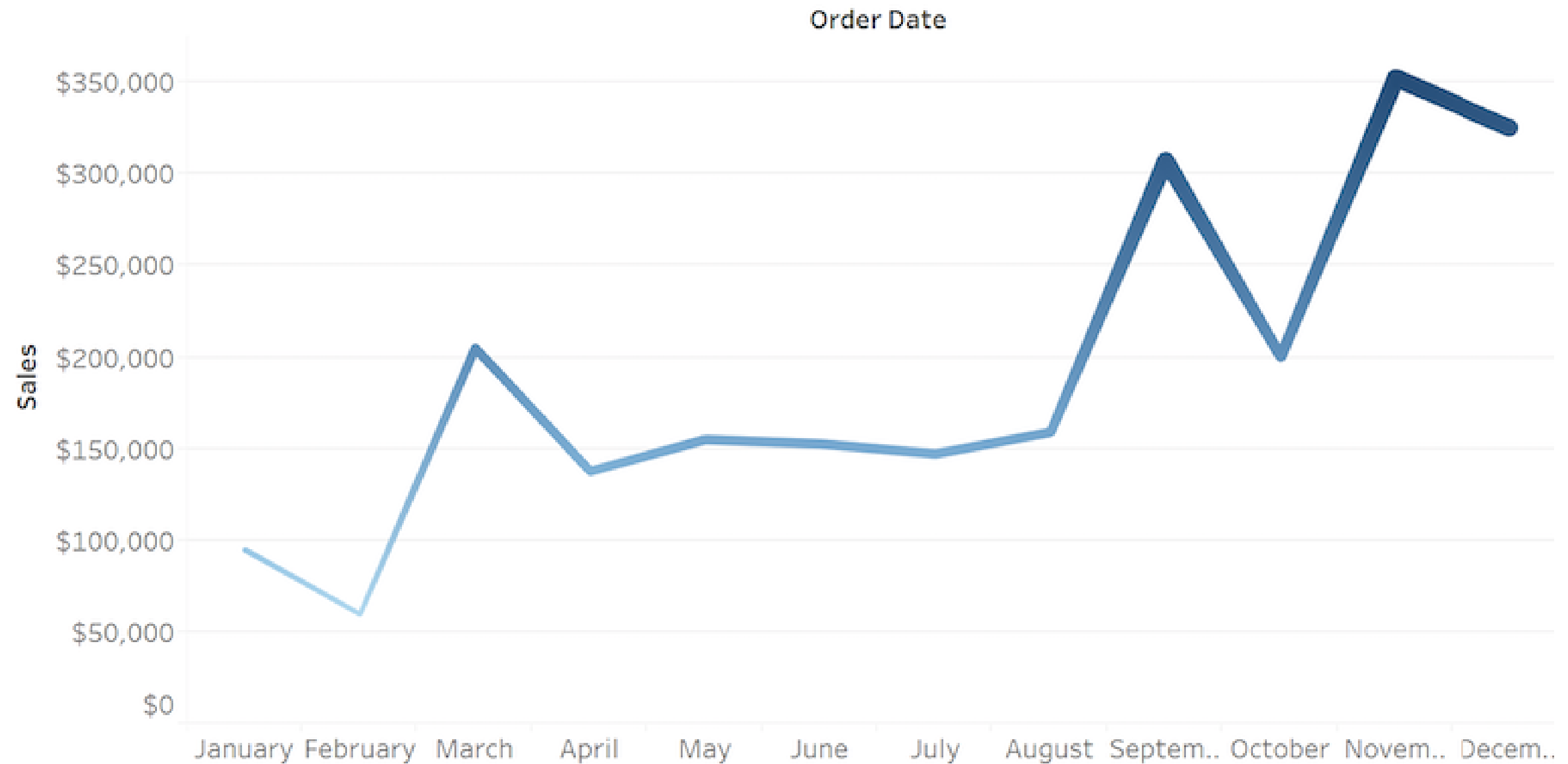
# Trends Over Time (Line Chart)



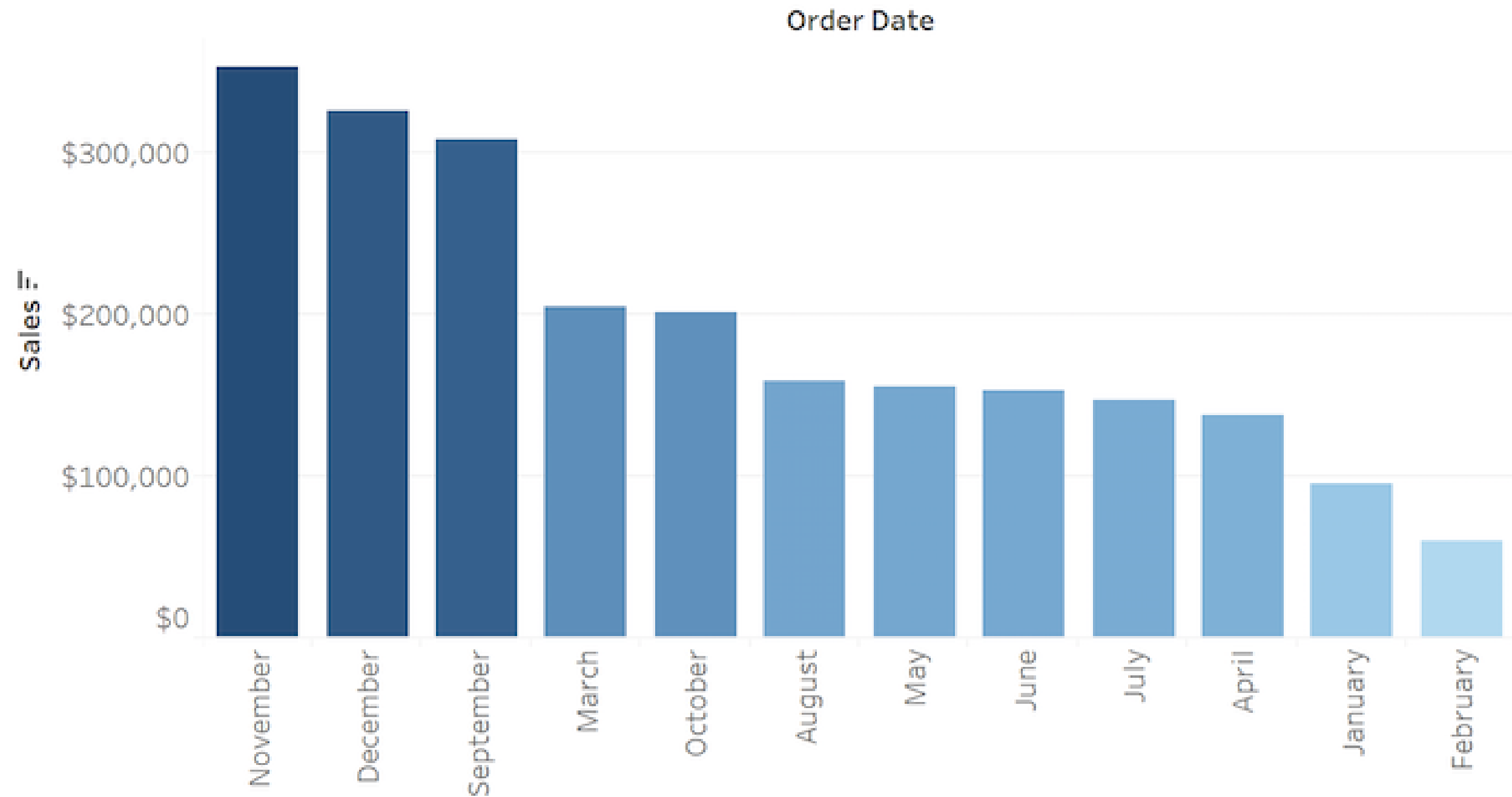
# Trends Over Time (Barchart)



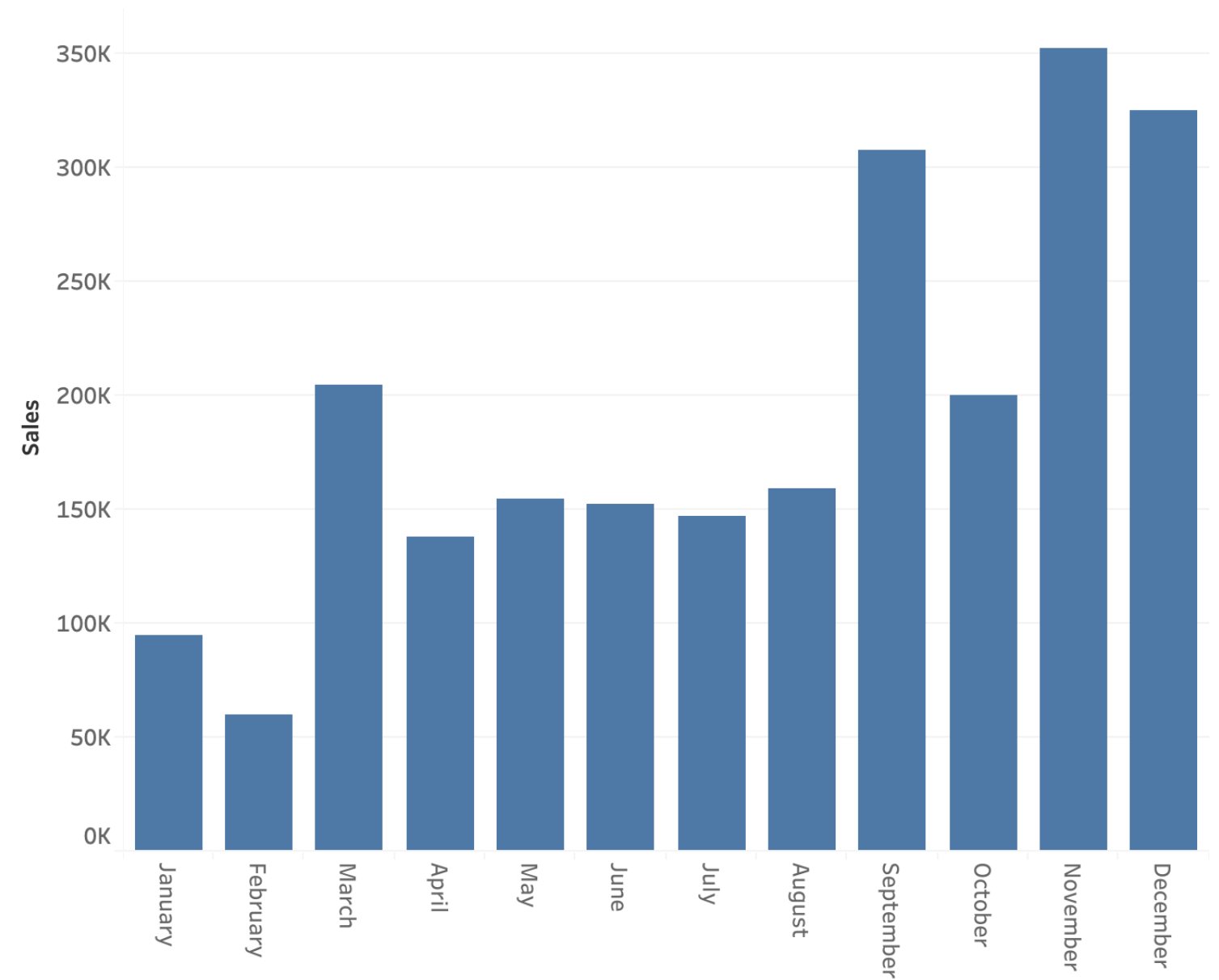
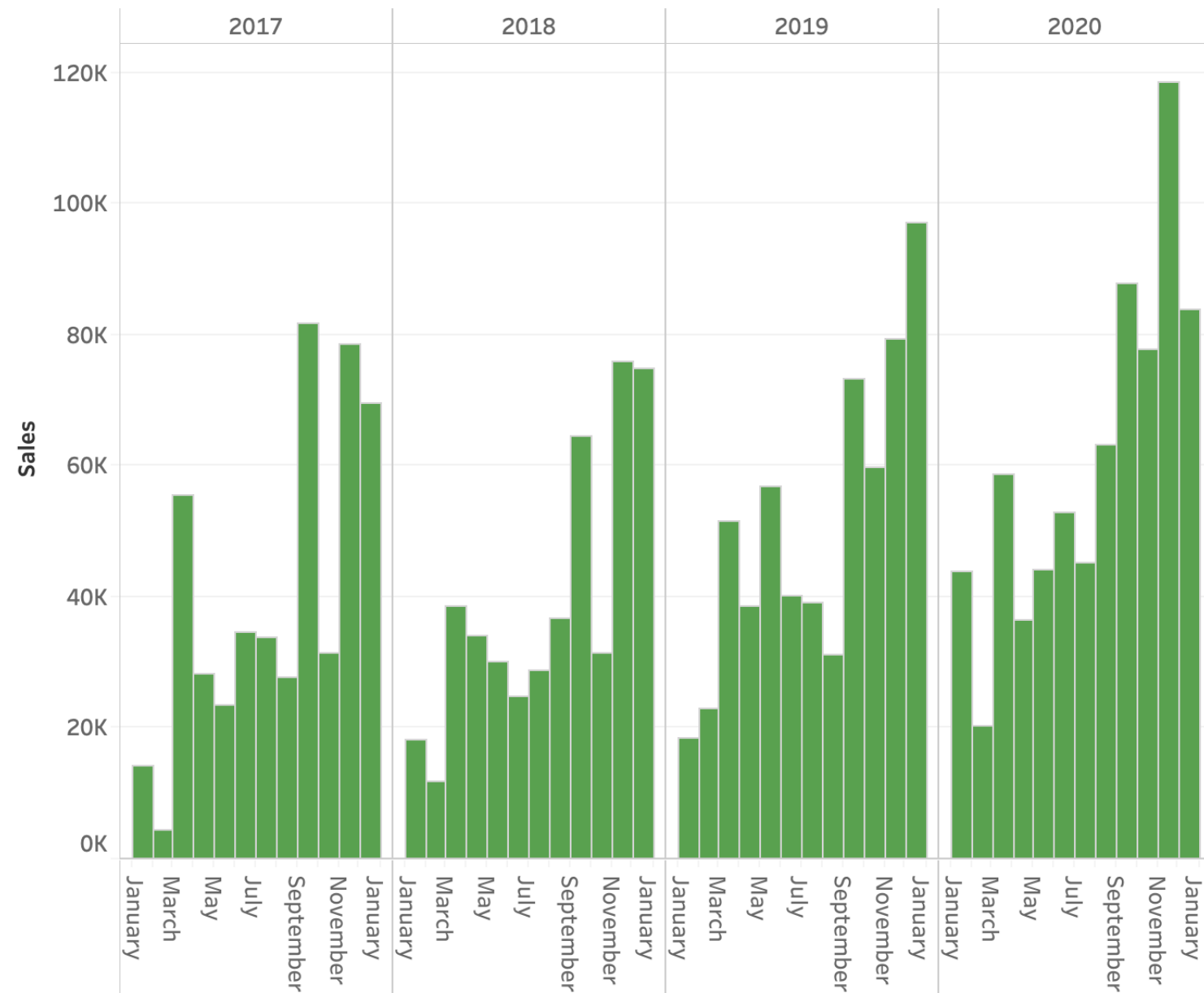
# Monthly Trendline of Sales Trends



# Monthly Pattern of Sales Trends



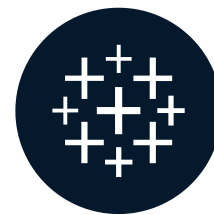
# Continuous vs. discrete



**Let's practice!**  
ANALYZING DATA IN TABLEAU

# Discrete time analysis and Quick Table Calculations

ANALYZING DATA IN TABLEAU



**Lis Sulmont**

Head of Curriculum at DataCamp

**Let's practice!**  
ANALYZING DATA IN TABLEAU



# Slicing and dicing

ANALYZING DATA IN TABLEAU



**Lis Sulmont**

Head of Curriculum at DataCamp

**Let's practice!**  
ANALYZING DATA IN TABLEAU