**What is data set**

## **What is a data set?**

A data set (sometimes referred to as data source, or database) in the context of Tableau, contains the data used to build visualizations. Every bar chart, scatter plot, or line chart you see in Tableau has a connected database or spreadsheet that supplies the data.

You can learn more about different data source types below.

Expand each **item**below to learn more about some data source types.

Spreadsheets

–

A spreadsheet, such as Microsoft Excel or Google sheets, organizes data in a flat structure, which means the records are stored as single rows of data.

Relational databases

–

Relational databases store data in multiple tables, with each row assigned a unique identifier. Users pull data from different tables together using Structured Query Language (SQL). The "relational" aspect indicates a logical connection between different tables.

Cloud data

–

Sometimes, organizations prefer to store their data in the cloud so they do not have to support on-premises servers. This includes data stored in such places as Amazon Web Services or Microsoft Azure.

Other types of connections

–

Tableau also connects to spatial files for mapping, such as .kml or .shp, and statistical files created in R.

# Supported Connectors

*Version: 2020.3  
Applies to: Tableau Desktop, Tableau Prep*

Follow the link below for information on how to connect to your specific data. Connectors are listed in the order that they appear on the **Connect**pane.

## Other articles in this section

* [Microsoft Excel](https://help.tableau.com/current/pro/desktop/en-us/examples_excel.htm)
* [Text File](https://help.tableau.com/current/pro/desktop/en-us/examples_text.htm)
* [Microsoft Access](https://help.tableau.com/current/pro/desktop/en-us/examples_access.htm)
* [JSON File](https://help.tableau.com/current/pro/desktop/en-us/examples_json.htm)
* [PDF File](https://help.tableau.com/current/pro/desktop/en-us/examples_pdf.htm)
* [Spatial File](https://help.tableau.com/current/pro/desktop/en-us/examples_spatial_files.htm)
* [Statistical File](https://help.tableau.com/current/pro/desktop/en-us/examples_statfile.htm)
* [Other Files](https://help.tableau.com/current/pro/desktop/en-us/examples_extract.htm)
* [Tableau Server or Tableau Online](https://help.tableau.com/current/pro/desktop/en-us/examples_tableauserver.htm)
* [Actian Matrix](https://help.tableau.com/current/pro/desktop/en-us/examples_actianmatrix.htm)
* [Actian Vectorwise](https://help.tableau.com/current/pro/desktop/en-us/examples_vectorwise.htm)
* [Alibaba AnalyticDB for MySQL](https://help.tableau.com/current/pro/desktop/en-us/examples_alibaba_analyticdb.htm)
* [Alibaba Data Lake Analytics](https://help.tableau.com/current/pro/desktop/en-us/examples_alibaba_data_lake_analytics.htm)
* [Alibaba MaxCompute](https://help.tableau.com/current/pro/desktop/en-us/examples_alibaba_maxcompute.htm)
* [Amazon Athena](https://help.tableau.com/current/pro/desktop/en-us/examples_amazonathena.htm)
* [Amazon Aurora for MySQL](https://help.tableau.com/current/pro/desktop/en-us/examples_amazonaurora.htm)
* [Amazon EMR Hadoop Hive](https://help.tableau.com/current/pro/desktop/en-us/examples_amazonemr.htm)
* [Amazon Redshift](https://help.tableau.com/current/pro/desktop/en-us/examples_amazonredshift.htm)
* [Anaplan](https://help.tableau.com/current/pro/desktop/en-us/examples_anaplan.htm)
* [Apache Drill](https://help.tableau.com/current/pro/desktop/en-us/examples_apachedrill.htm)
* [Aster Database](https://help.tableau.com/current/pro/desktop/en-us/examples_asterdata.htm)
* [Azure SQL Synapse Analytics](https://help.tableau.com/current/pro/desktop/en-us/examples_azure_sql_dw.htm)
* [Box](https://help.tableau.com/current/pro/desktop/en-us/examples_box.htm)
* [Cloudera Hadoop](https://help.tableau.com/current/pro/desktop/en-us/examples_hadoop.htm)
* [Databricks](https://help.tableau.com/current/pro/desktop/en-us/examples_databricks.htm)
* [Denodo](https://help.tableau.com/current/pro/desktop/en-us/examples_denodo.htm)
* [Dropbox](https://help.tableau.com/current/pro/desktop/en-us/examples_dropbox.htm)
* [Esri ArcGIS Server](https://help.tableau.com/current/pro/desktop/en-us/examples_esri.htm)
* [Exasol](https://help.tableau.com/current/pro/desktop/en-us/examples_exasolution.htm)
* [Firebird 3](https://help.tableau.com/current/pro/desktop/en-us/examples_firebird.htm)
* [Google Ads](https://help.tableau.com/current/pro/desktop/en-us/examples_googleads.htm)
* [Google Analytics](https://help.tableau.com/current/pro/desktop/en-us/examples_googleanalytics.htm)
* [Google BigQuery](https://help.tableau.com/current/pro/desktop/en-us/examples_googlebigquery.htm)
* [Google Cloud SQL](https://help.tableau.com/current/pro/desktop/en-us/examples_googlecloudsql.htm)
* [Google Drive](https://help.tableau.com/current/pro/desktop/en-us/examples_googledrive.htm)
* [Google Sheets](https://help.tableau.com/current/pro/desktop/en-us/examples_googlesheets.htm)
* [Hortonworks Hadoop Hive](https://help.tableau.com/current/pro/desktop/en-us/examples_hortonworkshadoop.htm)
* [IBM BigInsights](https://help.tableau.com/current/pro/desktop/en-us/examples_biginsights.htm)
* [IBM DB2](https://help.tableau.com/current/pro/desktop/en-us/examples_db2.htm)
* [IBM PDA (Netezza)](https://help.tableau.com/current/pro/desktop/en-us/examples_netezza.htm)
* [Impala](https://help.tableau.com/current/pro/desktop/en-us/examples_impala.htm)
* [Intuit QuickBooks Online](https://help.tableau.com/current/pro/desktop/en-us/examples_quickbooksonline.htm)
* [Kognitio](https://help.tableau.com/current/pro/desktop/en-us/examples_kognitio.htm)
* [Kyvos](https://help.tableau.com/current/pro/desktop/en-us/examples_kyvos.htm)
* [LinkedIn Sales Navigator](https://help.tableau.com/current/pro/desktop/en-us/examples_linkedin_sales_navigator.htm)
* [MapR Hadoop Hive](https://help.tableau.com/current/pro/desktop/en-us/examples_maprhadoop.htm)
* [MariaDB](https://help.tableau.com/current/pro/desktop/en-us/examples_mariadb.htm)
* [Marketo](https://help.tableau.com/current/pro/desktop/en-us/examples_marketo.htm)
* [MarkLogic](https://help.tableau.com/current/pro/desktop/en-us/examples_marklogic.htm)
* [MemSQL](https://help.tableau.com/current/pro/desktop/en-us/examples_memsql.htm)
* [Microsoft Analysis Services](https://help.tableau.com/current/pro/desktop/en-us/examples_msas.htm)
* [Microsoft PowerPivot](https://help.tableau.com/current/pro/desktop/en-us/examples_powerpivot.htm)
* [Microsoft SQL Server](https://help.tableau.com/current/pro/desktop/en-us/examples_sqlserver.htm)
* [MonetDB](https://help.tableau.com/current/pro/desktop/en-us/examples_monetdb.htm)
* [MongoDB BI Connector](https://help.tableau.com/current/pro/desktop/en-us/examples_mongodb.htm)
* [MySQL](https://help.tableau.com/current/pro/desktop/en-us/examples_mysql.htm)
* [OData](https://help.tableau.com/current/pro/desktop/en-us/examples_odata.htm)
* [OneDrive](https://help.tableau.com/current/pro/desktop/en-us/examples_onedrive.htm)
* [Oracle](https://help.tableau.com/current/pro/desktop/en-us/examples_oracle.htm)
* [Oracle Eloqua](https://help.tableau.com/current/pro/desktop/en-us/examples_eloqua.htm)
* [Oracle Essbase](https://help.tableau.com/current/pro/desktop/en-us/examples_essbase.htm)
* [Pivotal Greenplum](https://help.tableau.com/current/pro/desktop/en-us/examples_greenplum.htm)
* [PostgreSQL](https://help.tableau.com/current/pro/desktop/en-us/examples_postgresql.htm)
* [Presto](https://help.tableau.com/current/pro/desktop/en-us/examples_presto.htm)
* [Progress OpenEdge](https://help.tableau.com/current/pro/desktop/en-us/examples_progress.htm)
* [Qubole Presto](https://help.tableau.com/current/pro/desktop/en-us/examples_qubole.htm)
* [Salesforce](https://help.tableau.com/current/pro/desktop/en-us/examples_salesforce.htm)
* [Splunk](https://help.tableau.com/current/pro/desktop/en-us/examples_splunk.htm)
* [SAP HANA](https://help.tableau.com/current/pro/desktop/en-us/examples_saphana.htm)
* [SAP NetWeaver Business Warehouse](https://help.tableau.com/current/pro/desktop/en-us/examples_sapbw.htm)
* [SAP Sybase ASE](https://help.tableau.com/current/pro/desktop/en-us/examples_sybasease.htm)
* [SAP Sybase IQ](https://help.tableau.com/current/pro/desktop/en-us/examples_sybaseiq.htm)
* [ServiceNow ITSM](https://help.tableau.com/current/pro/desktop/en-us/examples_servicenow.htm)
* [SharePoint Lists](https://help.tableau.com/current/pro/desktop/en-us/examples_sharepoint_lists.htm)
* [Snowflake](https://help.tableau.com/current/pro/desktop/en-us/examples_snowflake.htm)
* [Spark SQL](https://help.tableau.com/current/pro/desktop/en-us/examples_sparksql.htm)
* [Teradata](https://help.tableau.com/current/pro/desktop/en-us/examples_teradata.htm)
* [Teradata OLAP Connector](https://help.tableau.com/current/pro/desktop/en-us/examples_teradata_olap.htm)
* [TIBCO Data Virtualization (Cisco Information Server)](https://help.tableau.com/current/pro/desktop/en-us/examples_ciscoinfoserver.htm)
* [Vertica](https://help.tableau.com/current/pro/desktop/en-us/examples_vertica.htm)
* [Connector Plugin](https://help.tableau.com/current/pro/desktop/en-us/examples_connector_sdk.htm)
* [Web Data Connector](https://help.tableau.com/current/pro/desktop/en-us/examples_web_data_connector.htm)
* [Other Databases (JDBC)](https://help.tableau.com/current/pro/desktop/en-us/examples_otherdatabases_jdbc.htm)
* [Other Databases (ODBC)](https://help.tableau.com/current/pro/desktop/en-us/examples_otherdatabases.htm)

### **What happens when Tableau Desktop connects to a data set?**

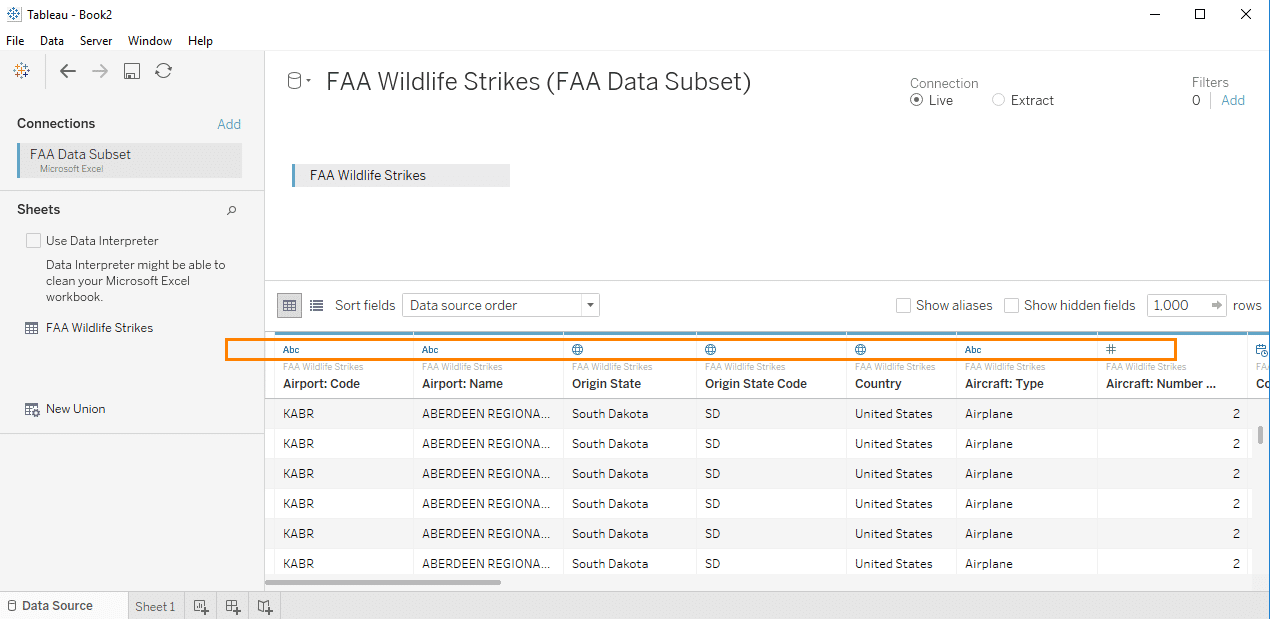
When users connect to Tableau, the data fields in their data set are automatically assigned a *role* and a *type*.

| **Role** | **Type** |
| --- | --- |
| A field can be assigned to a *Dimension*role or a *Measure*role.  We'll learn more about dimensions and measures in the next lesson. | The field's data type defines if the field is, for example, a string, integer, or date.  Data types can be changed by the user if Tableau does not assign the type appropriately. Any changes are saved in a Tableau data source (.tds) file as *metadata—*which is a set of data that describes other data*.* |

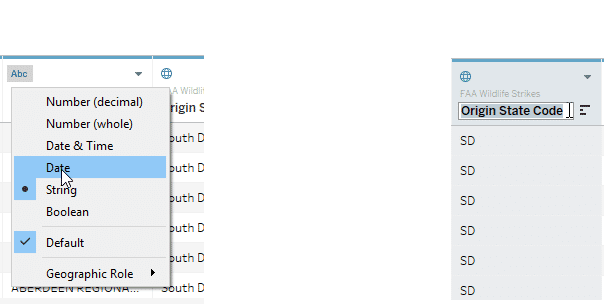
## **After Tableau connects to data...**

Click **Start**to view the general process after connection.

## **Connect**

After the connection, Tableau reads the data and assigns a data type to each field (see orange box).

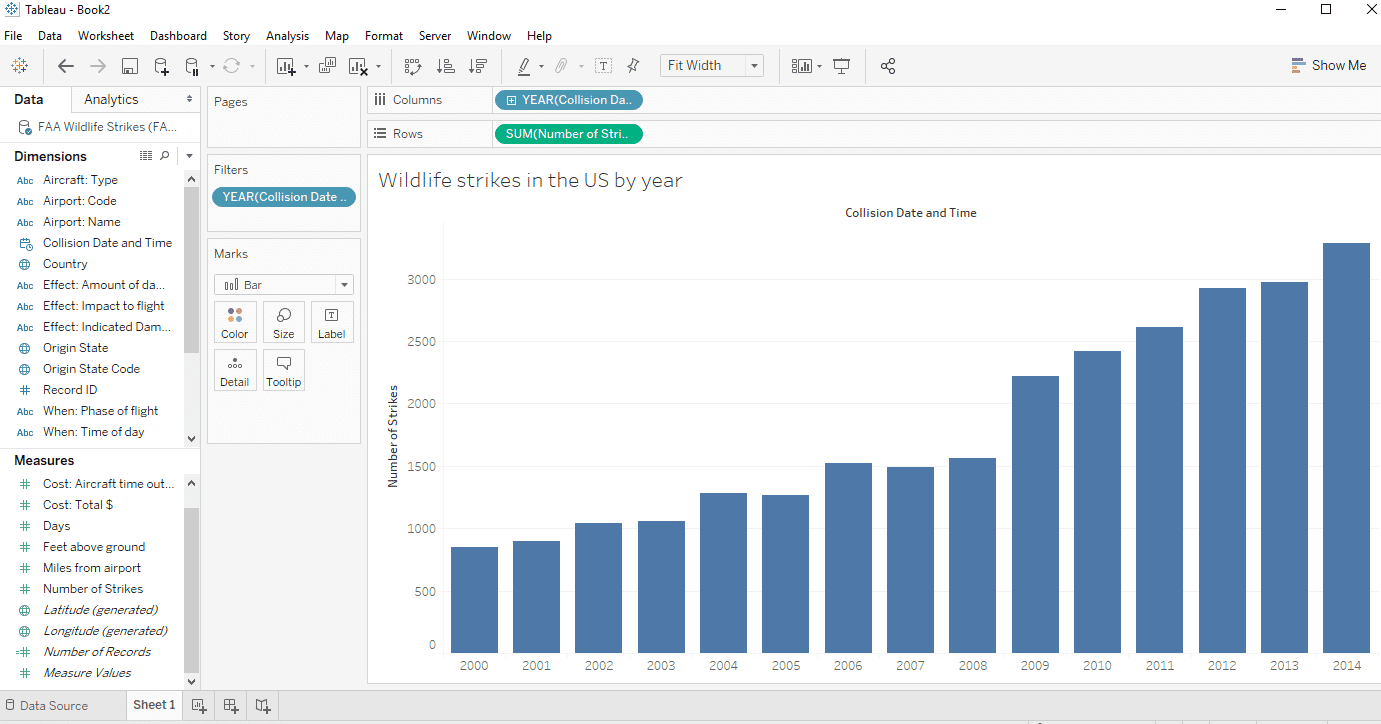
## **Evaluate**



Users verify that Tableau has correctly assigned data types, and they can make changes on the **Data Source** page if the data type is wrong. They can also rename unclear or misspelled fields.

Any changes made in Tableau are not written to the native data source. Changes are instead stored as metadata in a Tableau file called a Tableau Data Source, or .tds.

## **Analyze and Share**



After connecting to data and adjusting the metadata, users can open a sheet to begin building charts, analyzing data, and sharing results.

To read the chart you don’t need to be ab expert in database or spreadsheets but with a small amount of data knowledge you will feel more at ease in the world of data

**A closer look at fields**

**What is a field?**

A field, also known as a column, is a single piece of information from a record in a data set.

For example, if you were to collect data on how many times a week a commuter chose different modes of transportation, your data set could include:

| **Commuter name** | **Mode** | **Days per week** |
| --- | --- | --- |
| Neil | Walk | 2 |
| Neil | Bus | 3 |
| Lin | Bicycle | 1 |
| Lin | Carpool or Vanpool | 3 |
| Lin | Bus | 1 |
| Mae | Car (Single Occupancy Vehicle) | 2 |
| Mae | Bicycle | 3 |

*Commuter name*, *Mode*, and *Days per week* are all fields. Commuter name and Mode are qualitative fields, while Days per week is a quantitative field.

The members of the Commuter name and Mode field are limited to relevant categories. You would not include *apple*as an option in Mode because it doesn't match the definition of the field.

The Days per week uses a range of 1-7 rather than categorical members. If you wanted to collect data on modes that people did not use, your range might include 0, too.

You would not include negative values in the Days per week range, since you can't measure a negative trip. If you were measuring temperatures or profit, however, you could have a range that included negative numbers.

In Tableau, quantitative fields are referred to as *Measures*, and qualitative fields are referred to as *Dimensions*.

Flip each **card** below to learn more.

**Qualitative fields (Dimensions)**

Describes or categorizes data

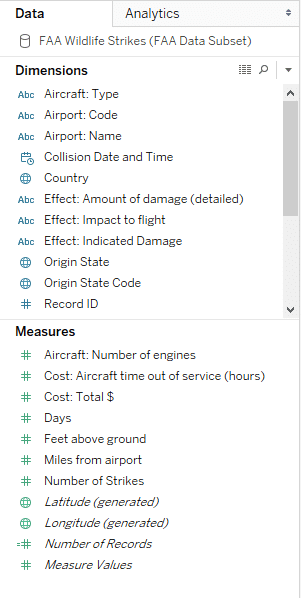
* Tells you what, when, or who
* Slices the quantitative data

**Quantitative fields (Measures)**

Numerical data

* Provides the measurement for qualitative category
* Can be used in calculations

**diamension and measures are the buildings blocks of tableau charts**

Tableau's **Data**pane, showing **Dimensions**and **Measures Fields in Tableau**

When Tableau connected to this data set, it assigned the fields to either Dimensions or Measures.

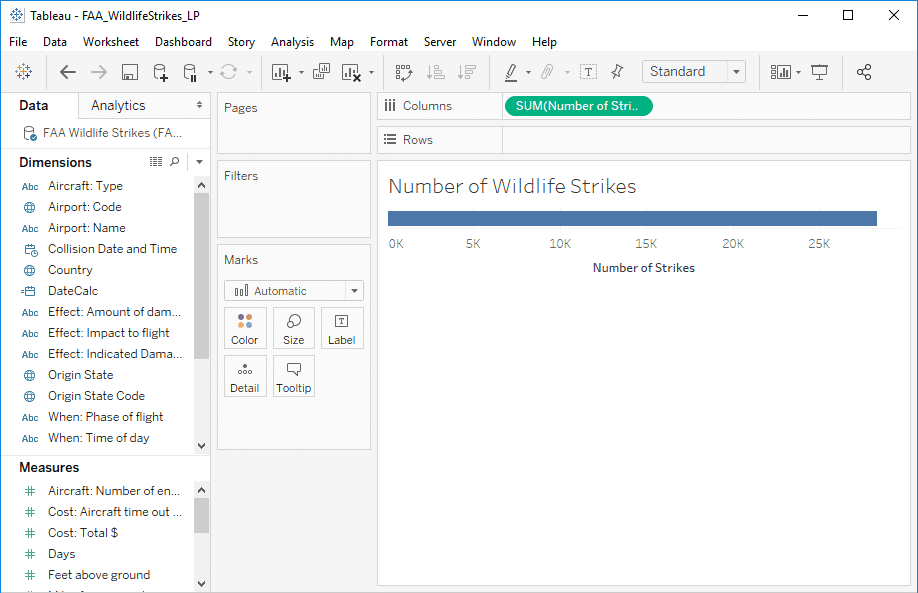
* The qualitative fields that describe categories of data are in the top part of the pane, under **Dimensions**.
* The quantitative fields that measure categories of data are in the bottom part of the pane, under **Measures**.

**How do these fields create a visualization?**

## **Let's build a chart!**

Even if you aren't a builder of charts, as a consumer you can better understand the finished product if you see how dimensions (qualitative) and measures (quantitative) are used to create visualizations and explore data.

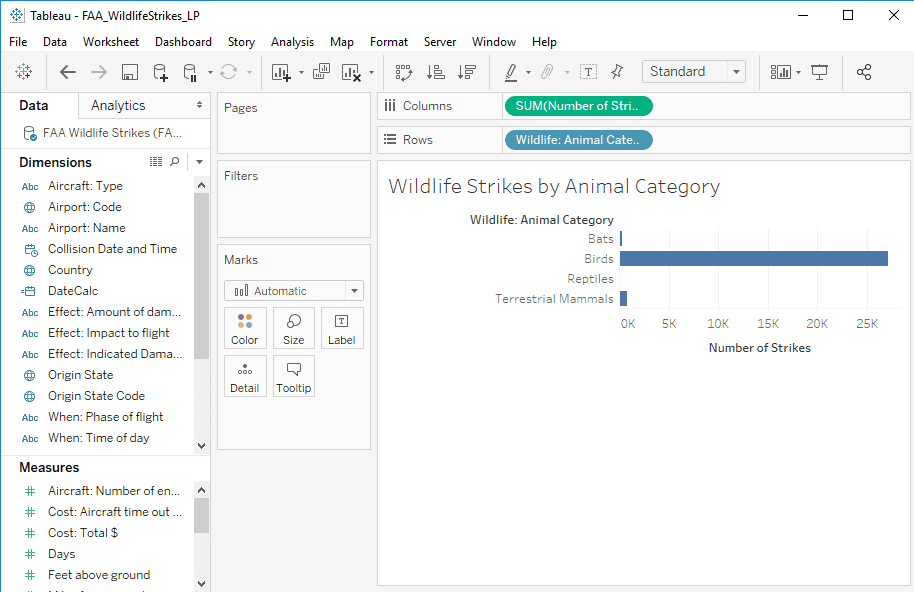
## **How many wildlife strikes occurred?**



First, drag out a quantitative field, or measure, to find out "how many?" We'll use **Number of Strikes**. Notice that the field displays on the **Columns** shelf in green.

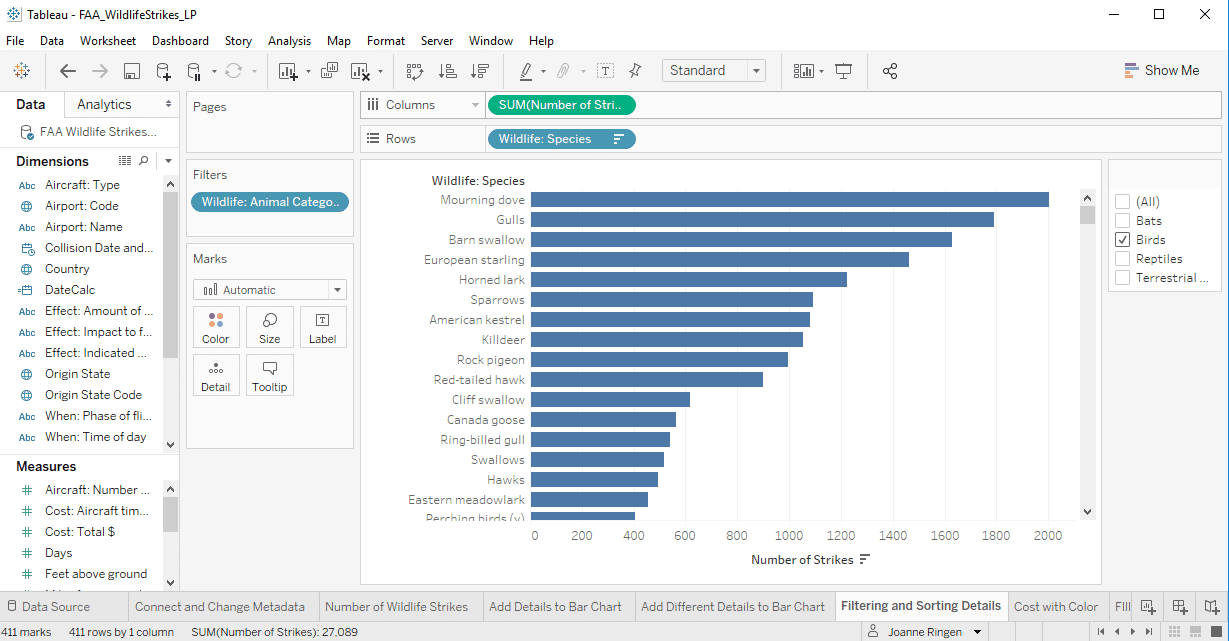
Tableau created a long bar and an axis showing a range of values.

## **How many wildlife strikes by animal category?**



We need to bring in qualitative fields, or dimensions, to better understand our data. If we add the dimension **Animal Category**, we can see that birds make up most of the wildlife strikes. The single bar breaks into four, one for each category. Dimension fields display in blue when brought onto the sheet.

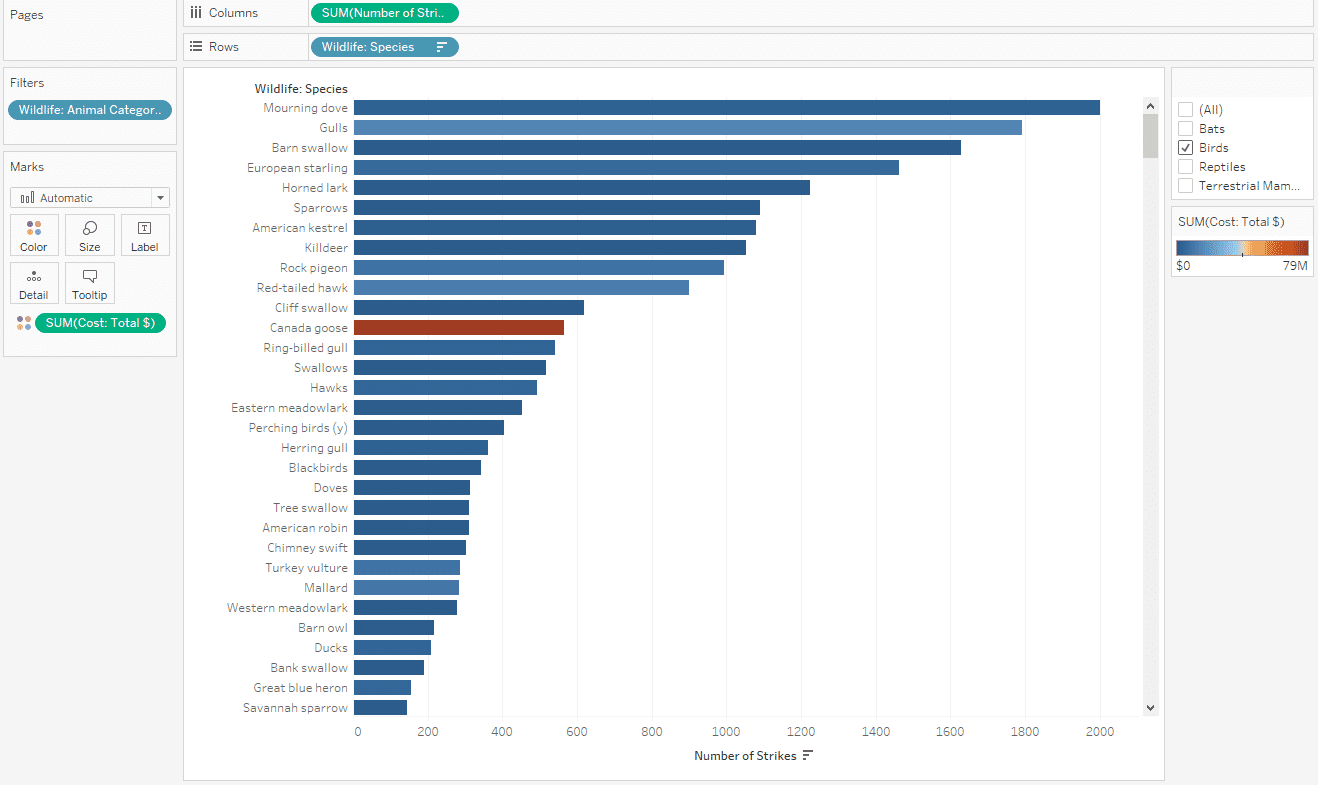
## **Which birds account for the most wildlife strikes?**



First, let's make **Animal Category** a filter, and add the **Animal Species** dimension field to look more closely at each category.

To see more details about the birds specifically, we filter the view to show only bird species.

## **Which species of bird is associated with the highest costs?**



We can add another measure field, **Cost: Total $**, to C**olor**to see which species of birds result in the costlies

t accidents.

It is interesting that the Canada goose is not involved in the highest number of bird strikes, but caused the most damage.

## **What did we discover?**

* As we added more dimensions to the view, the single bar representing all of the wildlife strikes was sliced into smaller categories of data.
* Using the filter allowed us to focus on a specific subset of data.
* We found an interesting data slice regarding the Canada goose that we should probably ask more questions about.

In Tableau, moving dimensions and measures in and out of the view changes the resulting chart. It's a useful way to view different aspects of your data, and helps uncover the trends and stories that affect your organization.

### **Do you want to learn more about fields in Tableau?**

Tableau's online help has more detailed information on fields and how they appear and behave in Tableau. See the link below for more information.