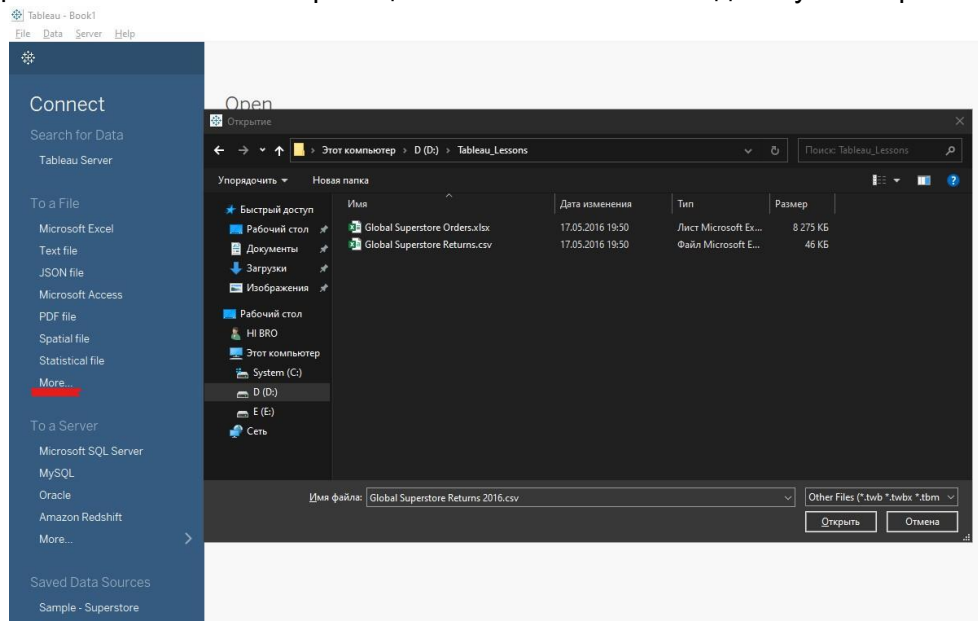


## Лабораторная работа №2. Визуализация данных в Tableau. Простые графики.

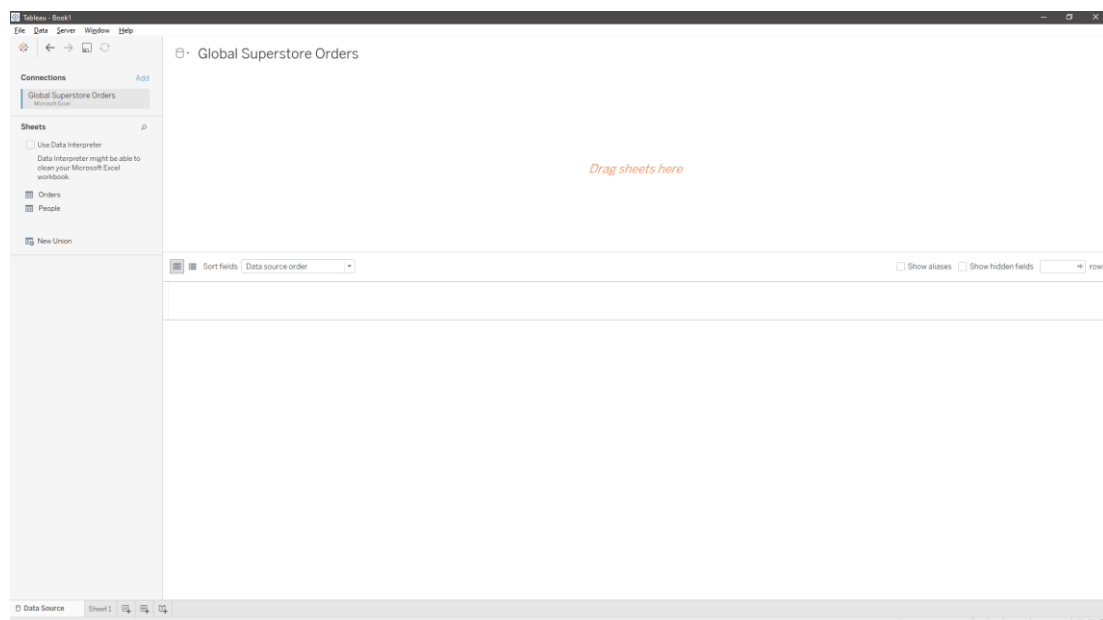
Устанавливаем программу Tableau с официального сайта. 2-ух недельного триала вполне хватит. Запускаем.

### Начало работы

Рассмотрим работу на примеры с таблицей покупок. Для начала откроем таблицу Excel. Для этого на главной странице жмем “More...” и находим нужный файл.



Откроется рабочий лист:



Файл Global Superstore Orders включает 2 таблицы (Заказы и Люди). Их можно увидеть слева на панели. Для отображения достаточно перетащить их на рабочую область (“Drag sheets here”).

Tableau - Book1

File Data Server Window Help

Connections: Global Superstore Orders (Microsoft Excel)

Sheets: Orders, People, New Union

Orders (Global Superstore Orders)

Sort fields: Data source order

Row ID	Alt Order ID	Alt Order Date	Alt Ship Date	Alt Ship Mode	Alt Customer ID	Alt Customer Name	Alt Segment	Alt Postal Code	Alt City	Alt State	Alt Country	Alt Region	Alt Market
40998	CA-2014-AB1001514...	11.11.2014	13.11.2014	First Class	AB-100151402	Aaron Bergman	Consumer	73120	Oklahoma City	Oklahoma	United States	Central US	USCA
26341	IN-2014-JR1621074...	05.02.2014	07.02.2014	Second Class	JR-162107	Justin Ritter	Corporate	null	Wollongong	New South Wales	Australia	Oceania	Asia Pacific
25330	IN-2014-CR1273074...	17.10.2014	18.10.2014	First Class	CR-127307	Craig Reiter	Consumer	null	Brisbane	Queensland	Australia	Oceania	Asia Pacific
13524	ES-2014-KM1637548...	28.01.2014	30.01.2014	First Class	KM-1637548	Katherine Murray	Home Office	null	Berlin	Berlin	Germany	Western Europe	Europe
47221	SG-2014-RH9495111...	05.11.2014	06.11.2014	Same Day	RH-9495111	Rick Hansen	Consumer	null	Dakar	Dakar	Senegal	Western Africa	Africa
22732	IN-2014-JM1565574...	28.06.2014	01.07.2014	Second Class	JM-156557	Jim Mitchum	Corporate	null	Sydney	New South Wales	Australia	Oceania	Asia Pacific
30570	IN-2013-TS2134092...	06.11.2012	08.11.2012	First Class	TS-2134092	Toby Swindell	Consumer	null	Porrus	Wellington	New Zealand	Oceania	Asia Pacific
31192	IN-2013-MB1808992...	14.04.2013	18.04.2013	Standard Class	MB-1808992	Mick Brown	Consumer	null	Hamilton	Waikato	New Zealand	Oceania	Asia Pacific
40099	CA-2014-AB1001514...	11.11.2014	13.11.2014	First Class	AB-100151402	Aaron Bergman	Consumer	73120	Oklahoma City	Oklahoma	United States	Central US	USCA
36258	CA-2012-AB1001514...	06.03.2012	07.03.2012	First Class	AB-100151404	Aaron Bergman	Consumer	98103	Seattle	Washington	United States	Western US	USCA
36259	CA-2012-AB1001514...	06.03.2012	07.03.2012	First Class	AB-100151404	Aaron Bergman	Consumer	98103	Seattle	Washington	United States	Western US	USCA
28879	ID-2013-AJ1078014...	19.04.2013	22.04.2013	First Class	AJ-107801	Anthony Jacobs	Corporate	null	Kabul	Kabul	Afghanistan	Southern Asia	Asia Pacific
45794	SA-2012-MM7260110...	26.12.2012	28.12.2012	Second Class	MM-7260110	Megdelene Morse	Consumer	null	Jizan	Jizan	Saudi Arabia	Western Asia	Asia Pacific
4132	MX-2013-VF2171518...	13.11.2013	13.11.2013	Same Day	VF-2171518	Vicky Freymann	Home Office	null	Toledo	Parana	Brazil	South America	LATAM
27704	IN-2014-PF1912027...	06.06.2014	08.06.2014	Second Class	PF-1912027	Peter Fuller	Consumer	null	Mudanjiang	Hellongjiang	China	Eastern Asia	Asia Pacific
13779	ES-2015-BP1118545...	31.07.2015	03.08.2015	Second Class	BP-1118545	Ben Peterman	Corporate	null	Paris	Ile-de-France	France	Western Europe	Europe

Go to Worksheet

Data Source

Если у нас есть связанные данные в другом источнике данных, можно создать интегрированный источник данных. Для этого надо добавить связь. Кликаем Добавить -> Текстовый Файл. Выберем имеющийся .csv файл.

Tableau - Book1

File Data Server Window Help

Connections: Global Superstore Orders (Microsoft Excel)

Sheets: Orders, People, New Union

Add a Connection

Search for Data

Tableau Server

To a File

Microsoft Excel

Text file

JSON file

Search

Action Matrix

Action Vector

Alibaba AnalyticDB for MySQL

Alibaba Data Lake Analytics

Alibaba MaxCompute

Amazon Athena

Amazon Aurora for MySQL

Amazon EMR Hadoop Hive

Открытие

Этот компьютер > D (D:) > Tableau\_Lessons

Глобальный поиск: Tableau\_Lessons

Имя	Дата изменения	Тип	Размер
Global Superstore Returns.csv	17.05.2016 19:50	Файл Microsoft E...	46 КБ

В результате получим таблицу с Join-ом. Как и в SQL можно выбрать тип join-a. Left join выдаст много значений null. (Если возникнут проблемы с подключением файлов, используйте файл initialized.twbx в папке с исходными данными).

Join dialog box configuration:

- Join Type: Inner
- Data Source: Global Superstore Returns 2...
- Field: Order ID
- Field: Order ID (Global Superst...)

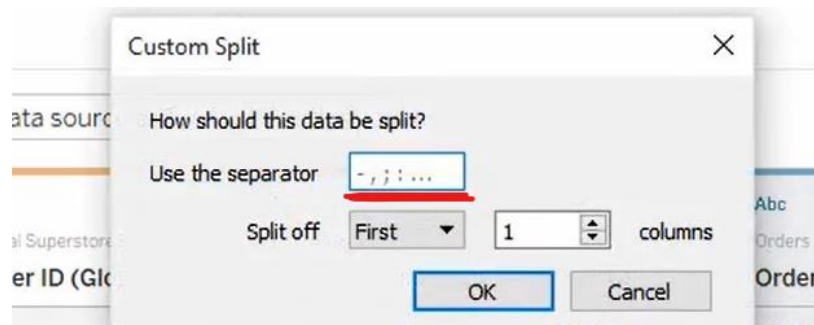
Returned	Order ID (Global S...	Region (Global Su...	Row ID	Order ID	Order Date	Ship D
Yes	IN-2012-PB19210127...	Eastern Asia	30191	IN-2012-PB19210127...	12/16/2012	12/19/
Yes	IN-2012-PB19210127...	Eastern Asia	30190	IN-2012-PB19210127...	12/16/2012	12/19/
Yes	IN-2015-JH158207-4...	Oceania	25438	IN-2015-JH158207-4...	5/16/2015	5/18/2
Yes	CA-2014-AS1004514...	Southern US	32648	CA-2014-AS1004514...	3/29/2014	4/1/20
Yes	IN-2014-LC168857-4...	Oceania	29629	IN-2014-LC168857-4...	4/18/2014	4/19/2
Yes	ID-2013-AB1001527-...	Eastern Asia	30267	ID-2013-AB1001527-...	6/14/2013	6/17/2
Yes	ES-2015-RA1994545-...	Western Europe	15162	ES-2015-RA1994545-...	8/2/2015	8/3/2
Yes	IN-2012-PR19210127	Eastern Asia	30187	IN-2012-PR19210127	12/16/2012	12/19/

Поле Order ID содержит в себе код центра доставки, год, код доставки. Такой вариант не всегда приятен для чтения, да и может не быть необходимости в такой избыточной информации. Для этого кликаем на «Дропдаун» в углу названия колонки и выбираем «Custom split»:

Dropdown menu options for 'Order ID':

- Rename
- Copy
- Hide
- Aliases...
- Create Calculated Field...
- Create Group...
- Split
- Custom Split...**
- Describe...

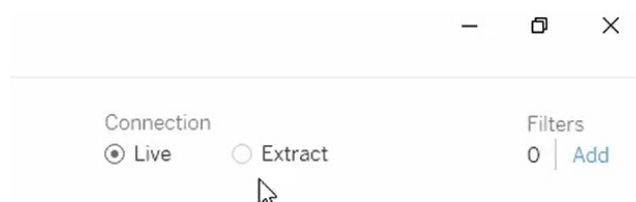
В поле разделителя поставим дефис и оставим первый элемент.



В итоге будет создана новая колонка только с кодами центров доставки.

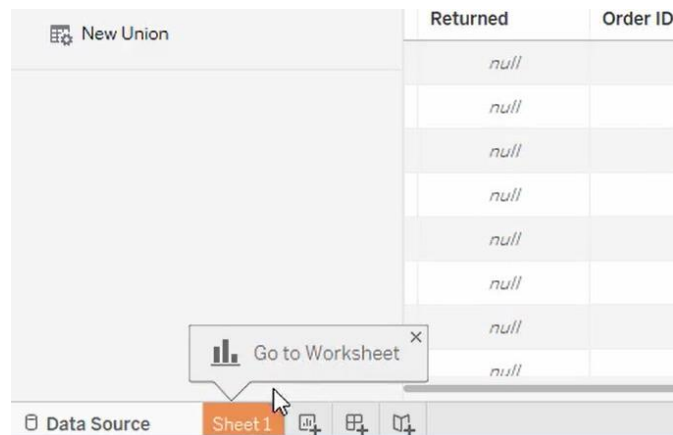
	Abc	=Abc	
	Orders	Calculation	Orders
D	Order ID	Order ID - Split 1	Order Date
3	CA-2014-AB10015140-41954	CA	11/11/2014
1	IN-2014-JR162107-41675	IN	2/5/2014
0	IN-2014-CR127307-41929	IN	10/17/2014
4	ES-2014-KM1637548-41667	ES	1/28/2014
1	SG-2014-RH9495111-41948	SG	11/5/2014
2	IN-2014-JM156557-41818	IN	6/28/2014

Два радио live и Extract отвечают за то, как будут читаться данные из источников. Live метод нужен для чтения быстроменяющихся в реальном времени данных, соответственно, требует больших затрат ресурсов. Extract подходит для случаев, когда необходимо провести анализ накопленных данных и требует меньших затрат.

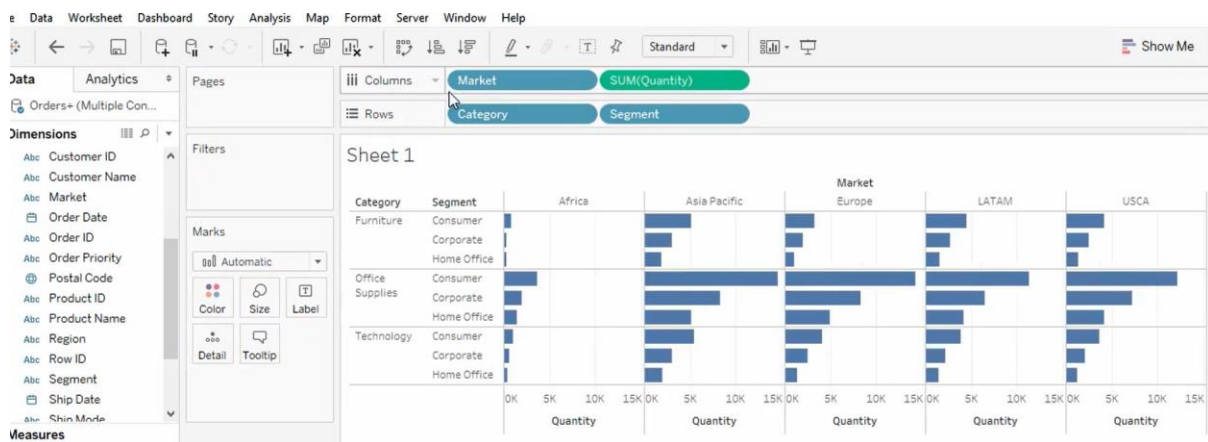


## Графики

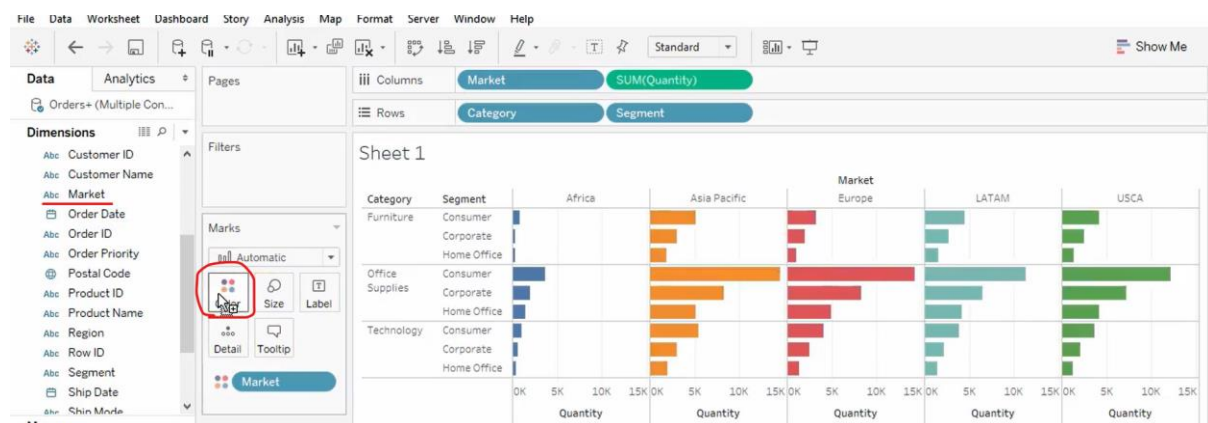
Теперь займемся обработкой данных. Для этого в левом нижнем «тулбаре» выберем наш лист (Лист 1). Откроется лист, на котором можно проводить различные манипуляции с нашими данными. Для начала визуализируем их. Для этого с левого «тулбара» перенесем поля Category и Segment в раздел «Строки», а Quantity и Market — в колоник.



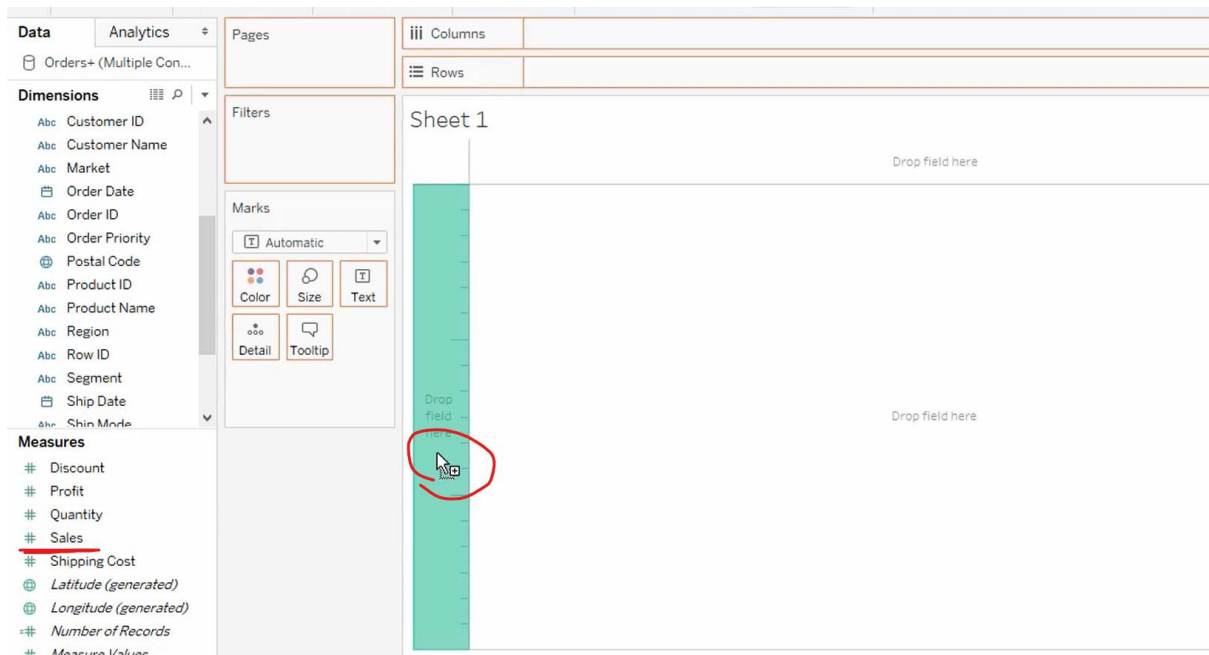
Получатся такие графики:



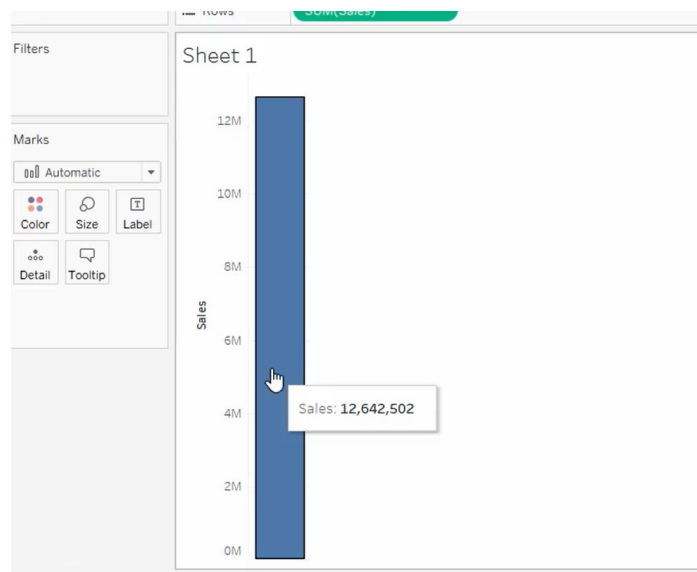
Для визуального улучшения внешнего вида добавим цветов. Для этого возьмем поле, отвечающее за каждый график (в данном случае Market), и перетащим его на «Colors». В итоге получим:



Теперь можно с полной уверенностью сказать, что Африка – не лучший рынок для нас. Теперь проанализируем отдельный показатель. Возьмем продажи с левой панели данных. И перетащим в область оси Y.

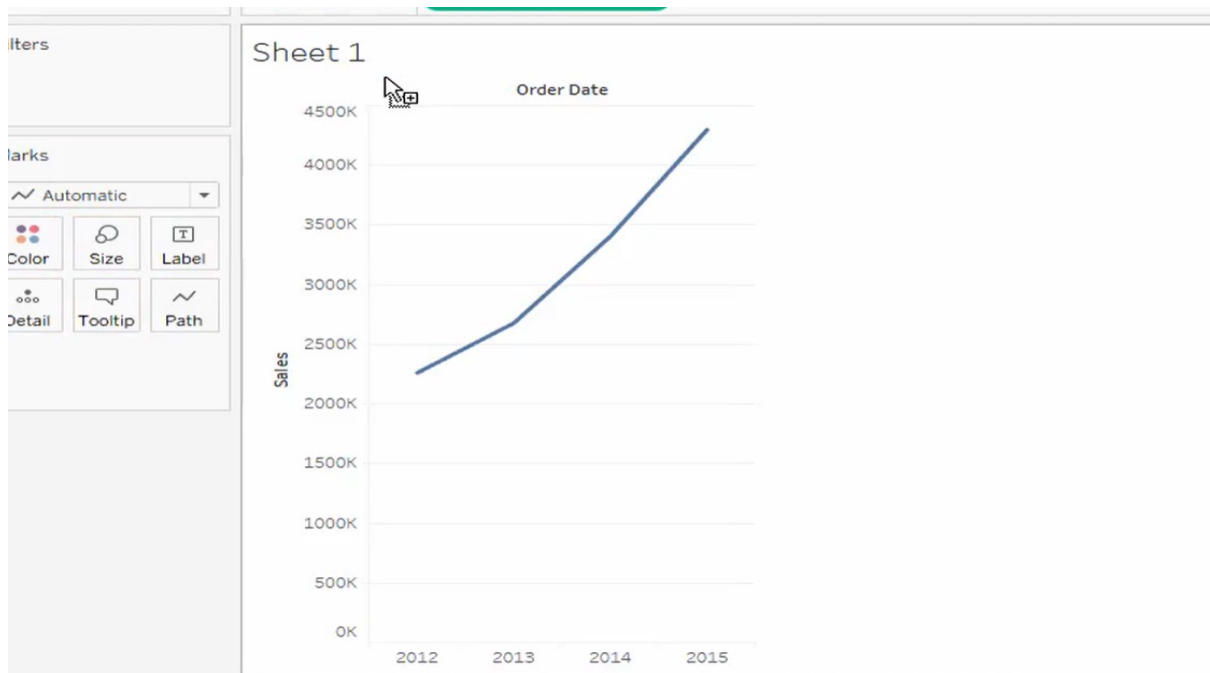


Сейчас мы видим общее количество продаж за все время.



Чтобы увидеть, как изменялись объемы продаж во времени, необходимо добавить вторым измерением поле даты (Order Date). Получим вот следующий график:

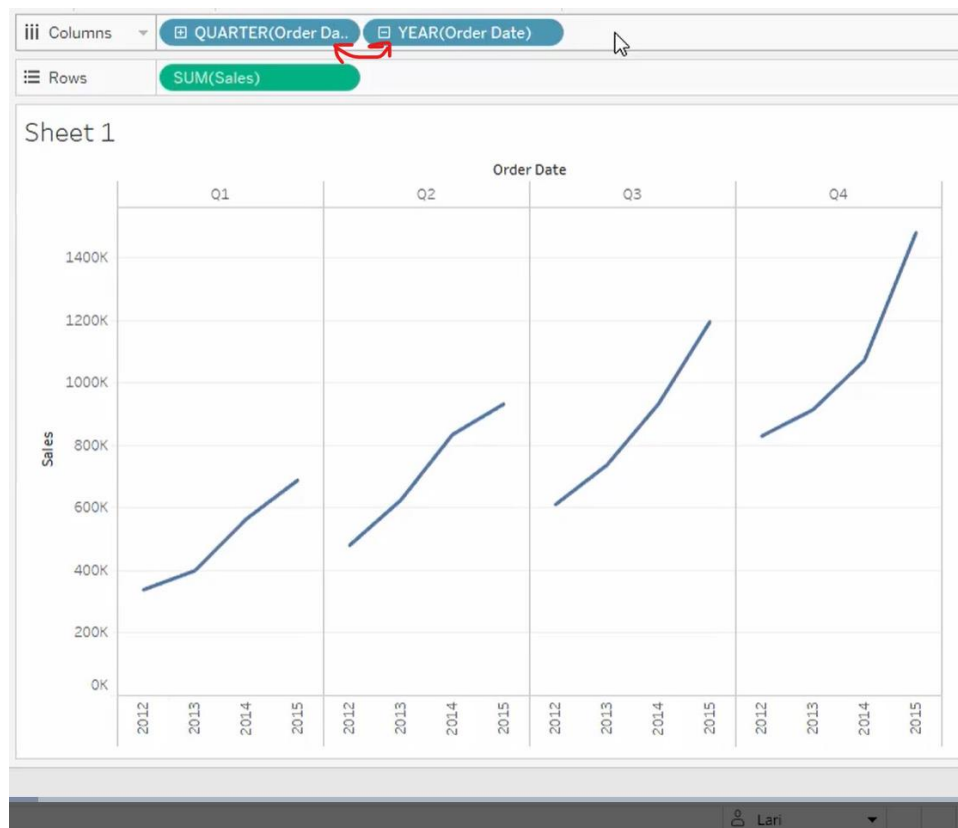




Если необходимо отразить данные за каждый год отдельно, в списке колонок можно нажать «+». Тогда будут построены графики для каждого года отдельно:



Если поменять местами Квартал и Год в поле колонок, то графики будут строиться относительно кварталов каждого года:

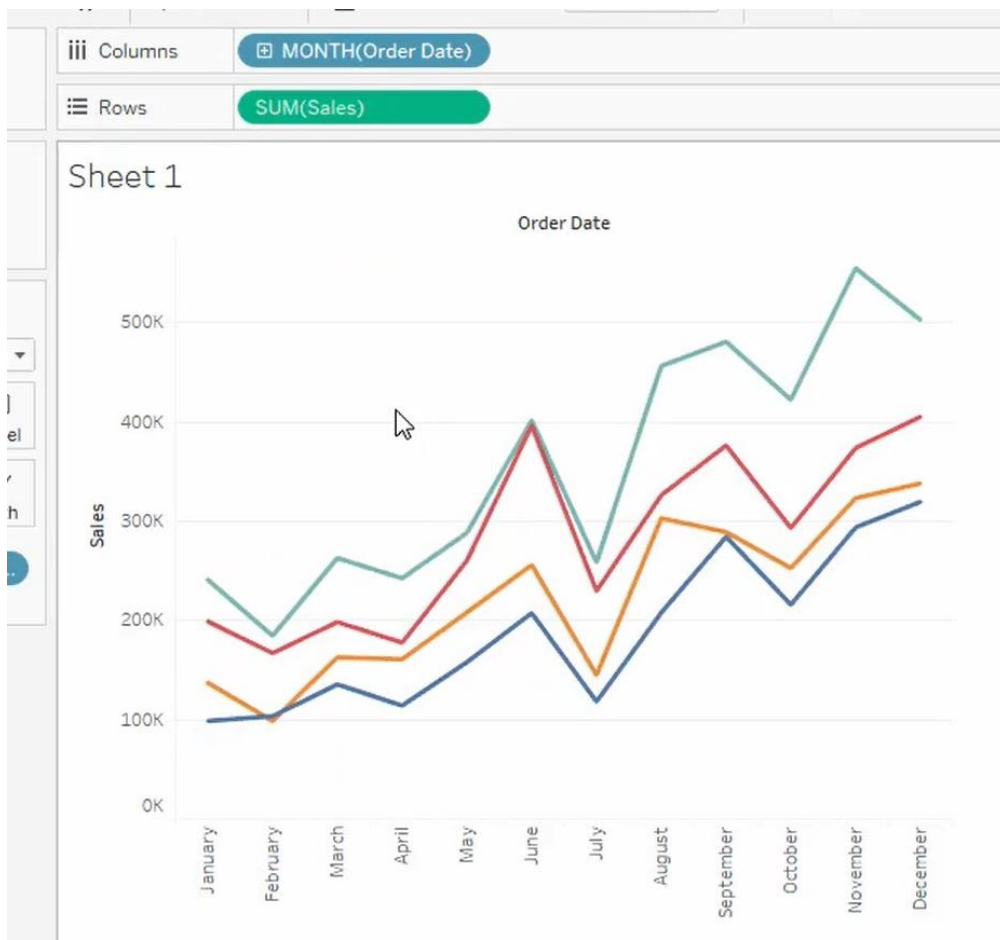


Возможностей визуализации тут немерено. Если убрать колонку года, будет следующая картина:

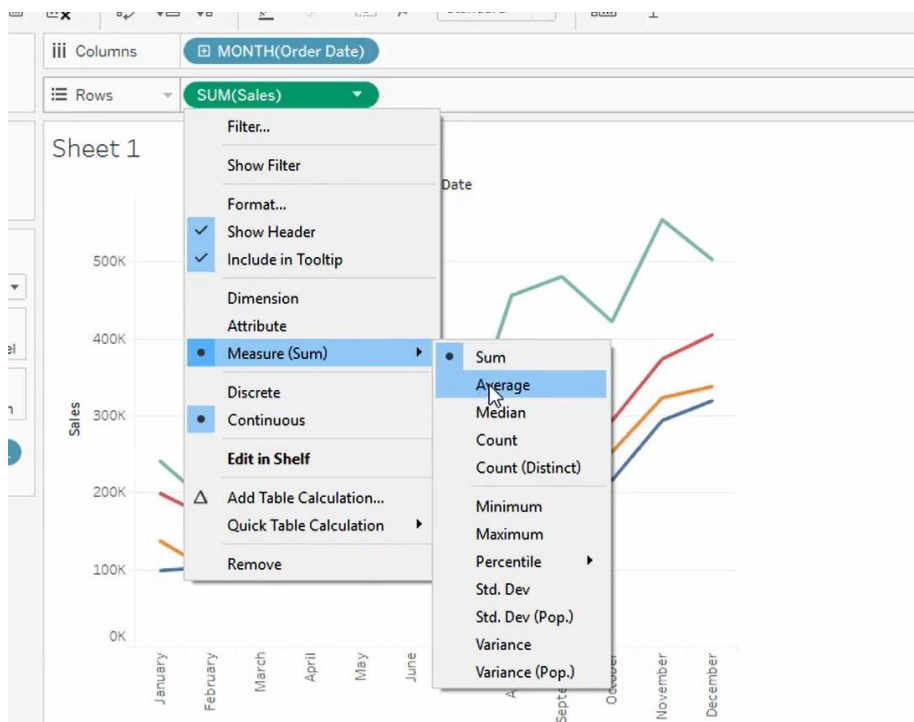


Если нажать на дропдаун оставшегося измерения (Month) и выбрать там промежутки вместо кварталов месяцы, то увидим следующий график.

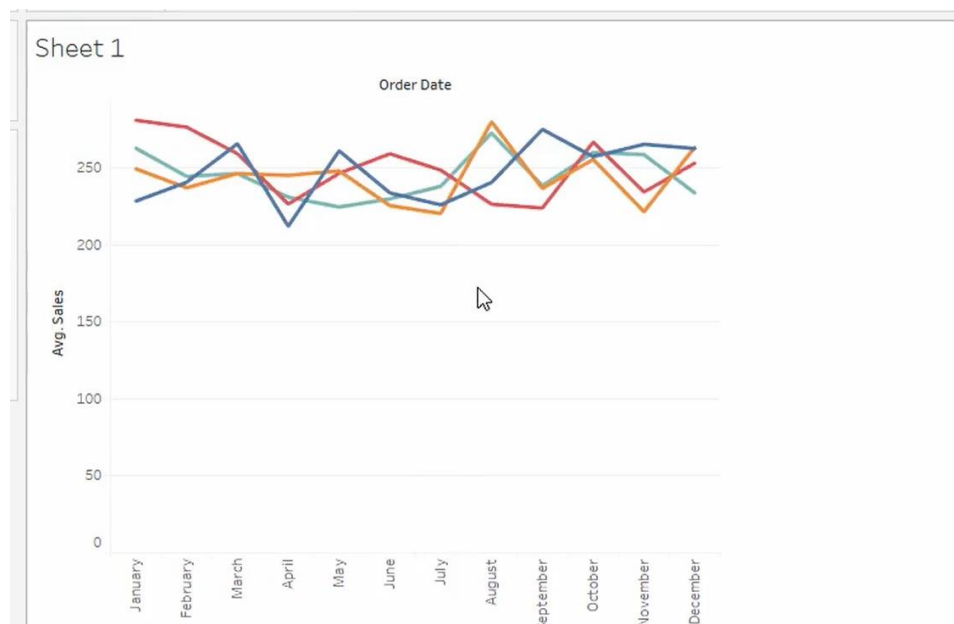




Для отображения среднего значения жмем на дропдаун поля Sales и вместо Sum выбираем Average:



В итоге получится такой график:



### Задание

Попрактикуйтесь в создании графиков по разным критериям. В следующей лабе будет интереснее. ; )