Module-03, Python for Data Analysis

Data Visualization (Seaborn)

Dostdar Ali
Instructor

Data science and Artificial Intelligence
3-Months Course
at
Karakaroum international Univrsity

January 7, 2024

Table of Contents

1. introduction to Seaborn
2. Distribution Plots
3. Categorical Data Plots
4. Matrix Plots
5. Regression Plots

Seaborn

-
- Data, Seaborn comes with built-in data sets!

	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4

Seaborn

-
- Data, Seaborn comes with built-in data sets!

	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4

Distribution Plots

- Such plots, allow us to visualize the distribution of a data set.
- These plots are: distplot, jointplot, pairplot, rugplot, kdeplot.
- distplot
  The distplot shows the distribution of a univariate set of observations.
- jointplot
  jointplot() allows you to basically match up two distplots for bivariate data. With your choice of what kind parameter to compare with:
- pairplot
  pairplot will plot pairwise relationships across an entire dataframe (for the numerical columns) and supports a color hue argument (for categorical columns).
- rugplot
  rugplots are actually a very simple concept, they just draw a dash mark for every point on a univariate distribution. They are the building block of a KDE plot:

Distribution Plots

- Such plots, allow us to visualize the distribution of a data set.
- These plots are: distplot, jointplot, pairplot, rugplot, kdeplot.
- distplot
  The distplot shows the distribution of a univariate set of observations.
- jointplot
  jointplot() allows you to basically match up two distplots for bivariate data. With your choice of what kind parameter to compare with:
- pairplot
  pairplot will plot pairwise relationships across an entire dataframe (for the numerical columns) and supports a color hue argument (for categorical columns).
- rugplot
  rugplots are actually a very simple concept, they just draw a dash mark for every point on a univariate distribution. They are the building block of a KDE plot:

Distribution Plots

- Such plots, allow us to visualize the distribution of a data set.
- These plots are: distplot, jointplot, pairplot, rugplot, kdeplot.
- distplot
  The distplot shows the distribution of a univariate set of observations.
- jointplot
  jointplot() allows you to basically match up two distplots for bivariate data. With your choice of what kind parameter to compare with:
- pairplot
  pairplot will plot pairwise relationships across an entire dataframe (for the numerical columns) and supports a color hue argument (for categorical columns).
- rugplot
  rugplots are actually a very simple concept, they just draw a dash mark for every point on a univariate distribution. They are the building block of a KDE plot:

Distribution Plots

- Such plots, allow us to visualize the distribution of a data set.
- These plots are: distplot, jointplot, pairplot, rugplot, kdeplot.
- distplot
  The distplot shows the distribution of a univariate set of observations.
- jointplot
  jointplot() allows you to basically match up two distplots for bivariate data. With your choice of what kind parameter to compare with:
- pairplot
  pairplot will plot pairwise relationships across an entire dataframe (for the numerical columns) and supports a color hue argument (for categorical columns).
- rugplot
  rugplots are actually a very simple concept, they just draw a dash mark for every point on a univariate distribution. They are the building block of a KDE plot:

Distribution Plots

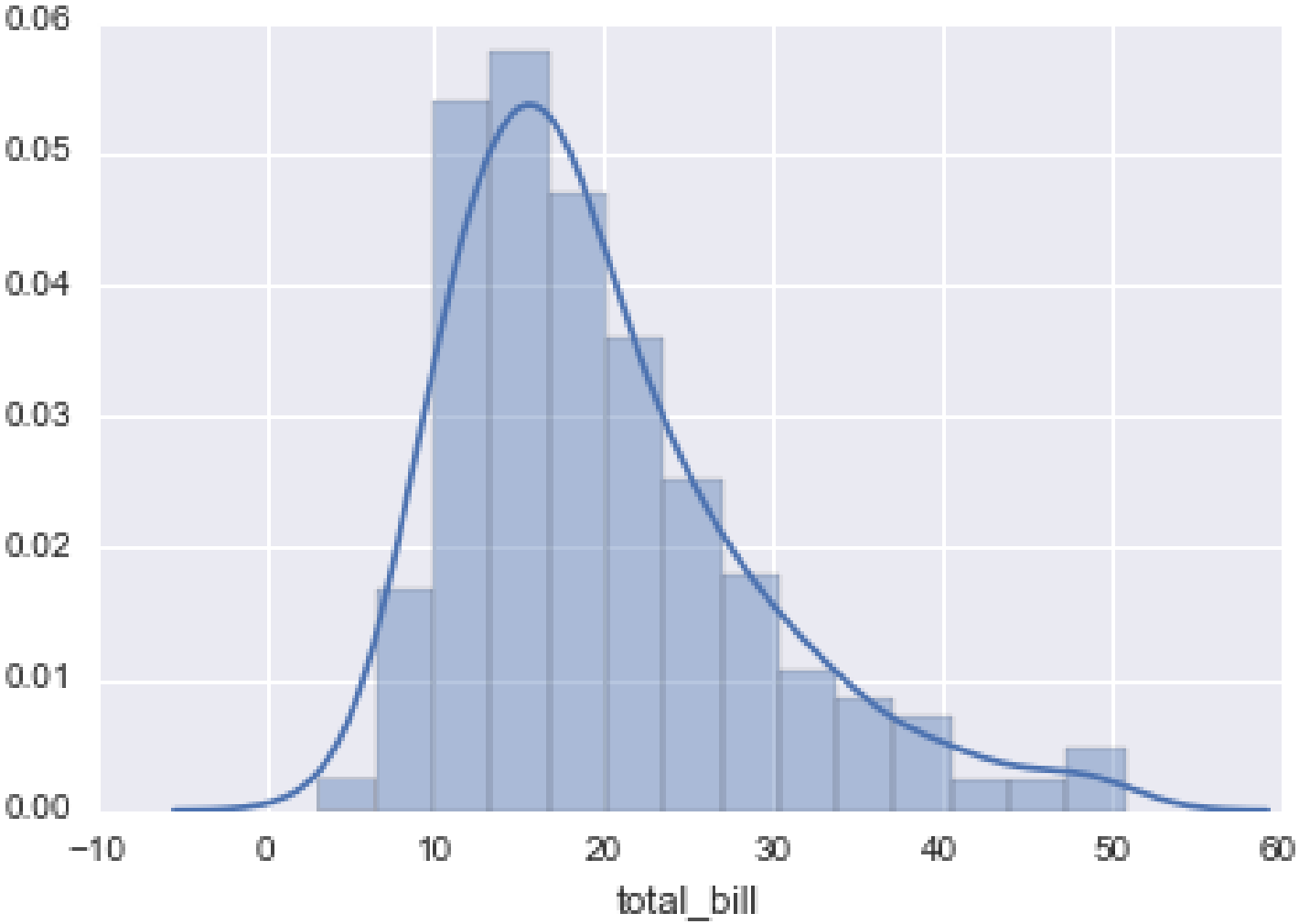
- Such plots, allow us to visualize the distribution of a data set.
- These plots are: distplot, jointplot, pairplot, rugplot, kdeplot.
- distplot
  The distplot shows the distribution of a univariate set of observations.
- jointplot
  jointplot() allows you to basically match up two distplots for bivariate data. With your choice of what kind parameter to compare with:
- pairplot
  pairplot will plot pairwise relationships across an entire dataframe (for the numerical columns) and supports a color hue argument (for categorical columns).
- rugplot
  rugplots are actually a very simple concept, they just draw a dash mark for every point on a univariate distribution. They are the building block of a KDE plot:

Distribution Plots

- Such plots, allow us to visualize the distribution of a data set.
- These plots are: distplot, jointplot, pairplot, rugplot, kdeplot.
- distplot
  The distplot shows the distribution of a univariate set of observations.
- jointplot
  jointplot() allows you to basically match up two distplots for bivariate data. With your choice of what kind parameter to compare with:
- pairplot
  pairplot will plot pairwise relationships across an entire dataframe (for the numerical columns) and supports a color hue argument (for categorical columns).
- rugplot
  rugplots are actually a very simple concept, they just draw a dash mark for every point on a univariate distribution. They are the building block of a KDE plot:

Distribution plot

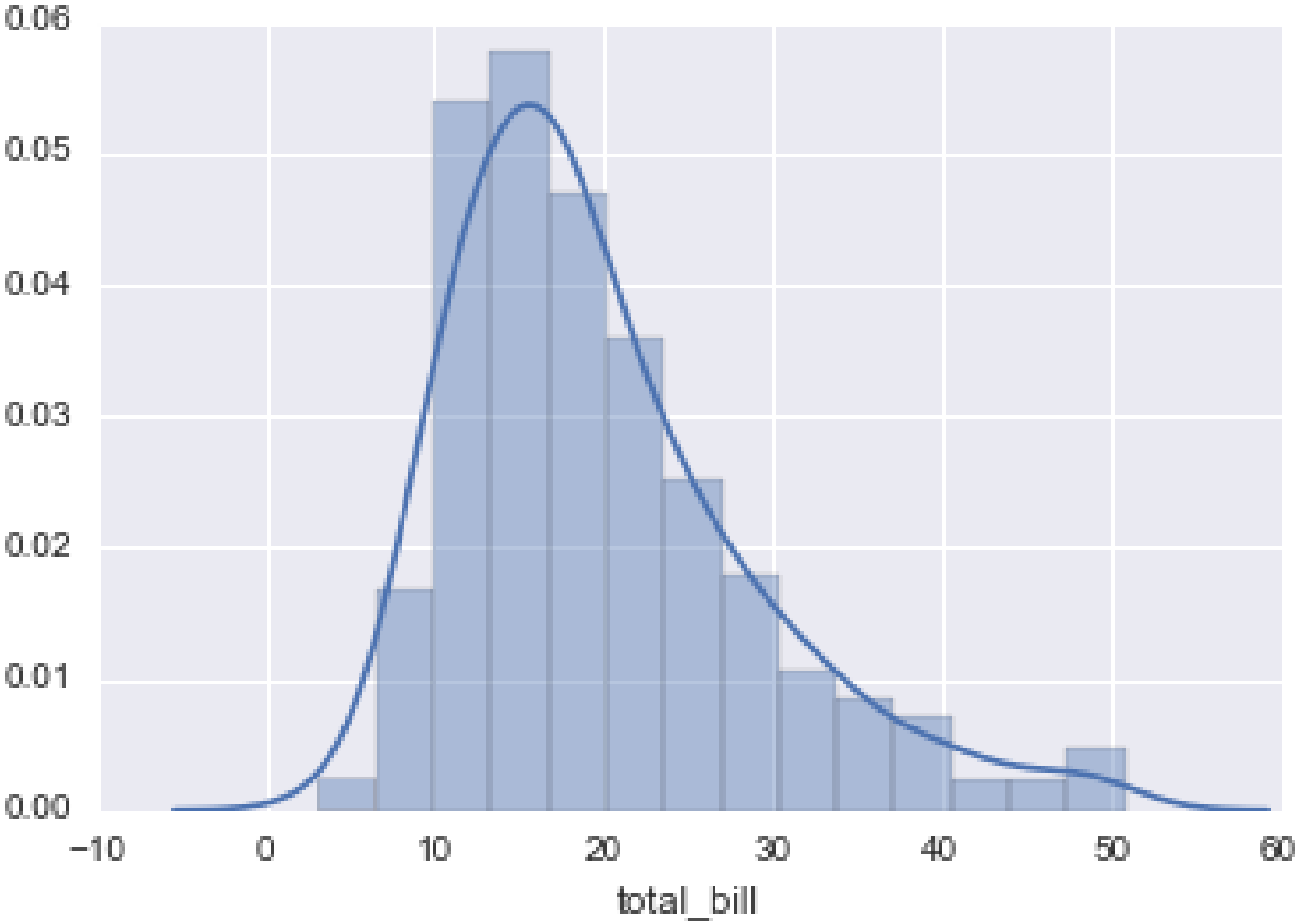
- Distribution Plot,



- Totall bill.

Distribution plot

- Distribution Plot,



- Totall bill.

Categorical Data Plots

Seaborn to plot categorical data! There are a few main plot types for this:

- factorplot
- boxplot
- violinplot
- stripplot
- swarmplot
- barplot
- countplot

Categorical Data Plots

Seaborn to plot categorical data! There are a few main plot types for this:

- factorplot
- boxplot
- violinplot
- stripplot
- swarmplot
- barplot
- countplot

Categorical Data Plots

Seaborn to plot categorical data! There are a few main plot types for this:

- factorplot
- boxplot
- violinplot
- stripplot
- swarmplot
- barplot
- countplot

Categorical Data Plots

Seaborn to plot categorical data! There are a few main plot types for this:

- factorplot
- boxplot
- violinplot
- stripplot
- swarmplot
- barplot
- countplot

Categorical Data Plots

Seaborn to plot categorical data! There are a few main plot types for this:

- factorplot
- boxplot
- violinplot
- stripplot
- swarmplot
- barplot
- countplot

Categorical Data Plots

Seaborn to plot categorical data! There are a few main plot types for this:

- factorplot
- boxplot
- violinplot
- stripplot
- swarmplot
- barplot
- countplot

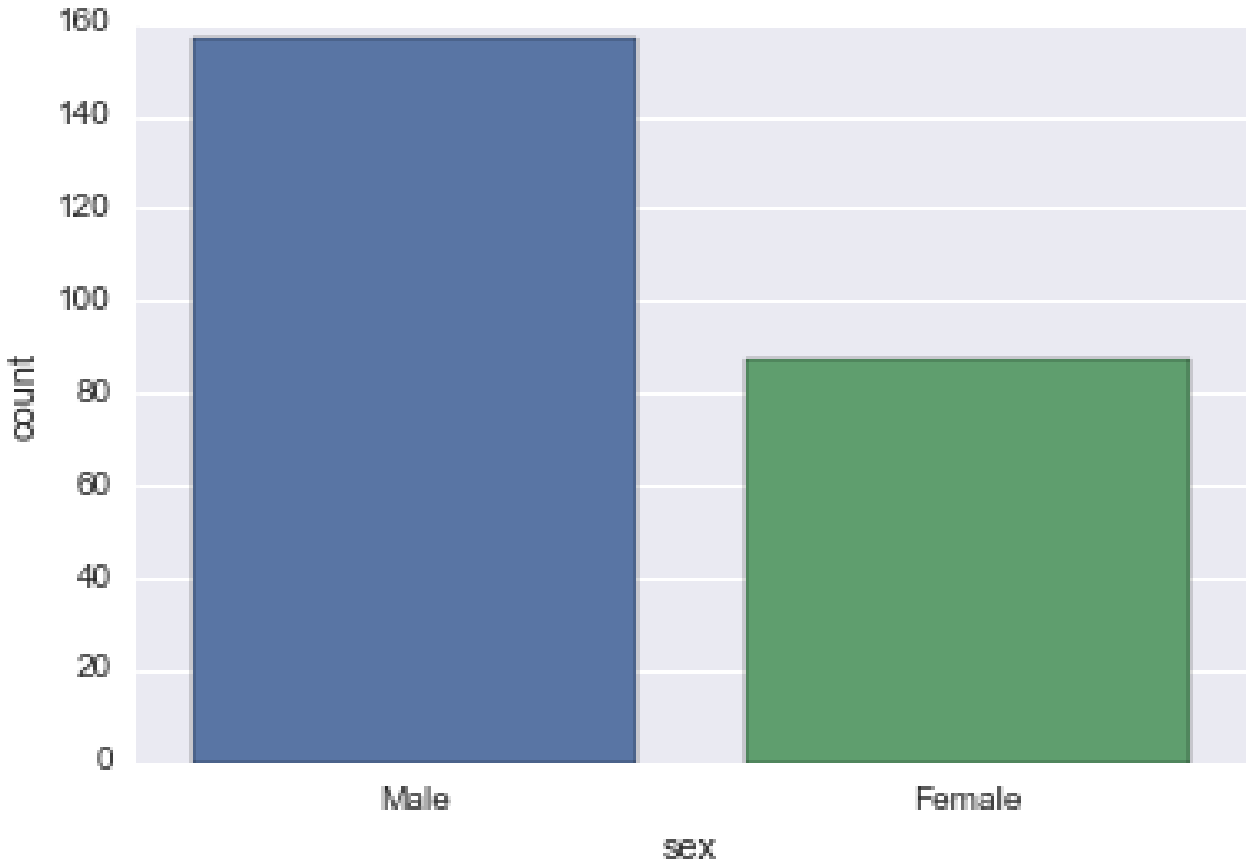
Categorical Data Plots

Seaborn to plot categorical data! There are a few main plot types for this:

- factorplot
- boxplot
- violinplot
- stripplot
- swarmplot
- barplot
- countplot

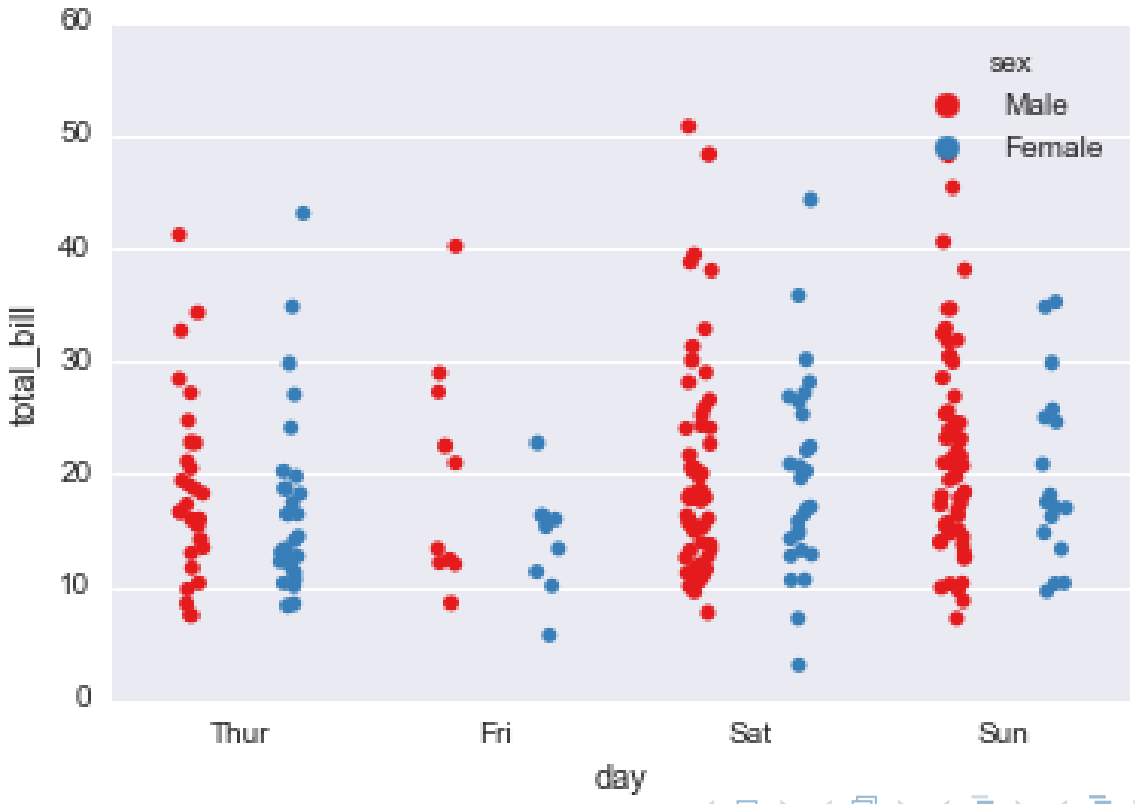
barplot and countplot

- These very similar plots allow you to get aggregate data off a categorical feature in your data. barplot is a general plot that allows you to aggregate the categorical data based off some function, by default the mean,



stripplot and swarmplot

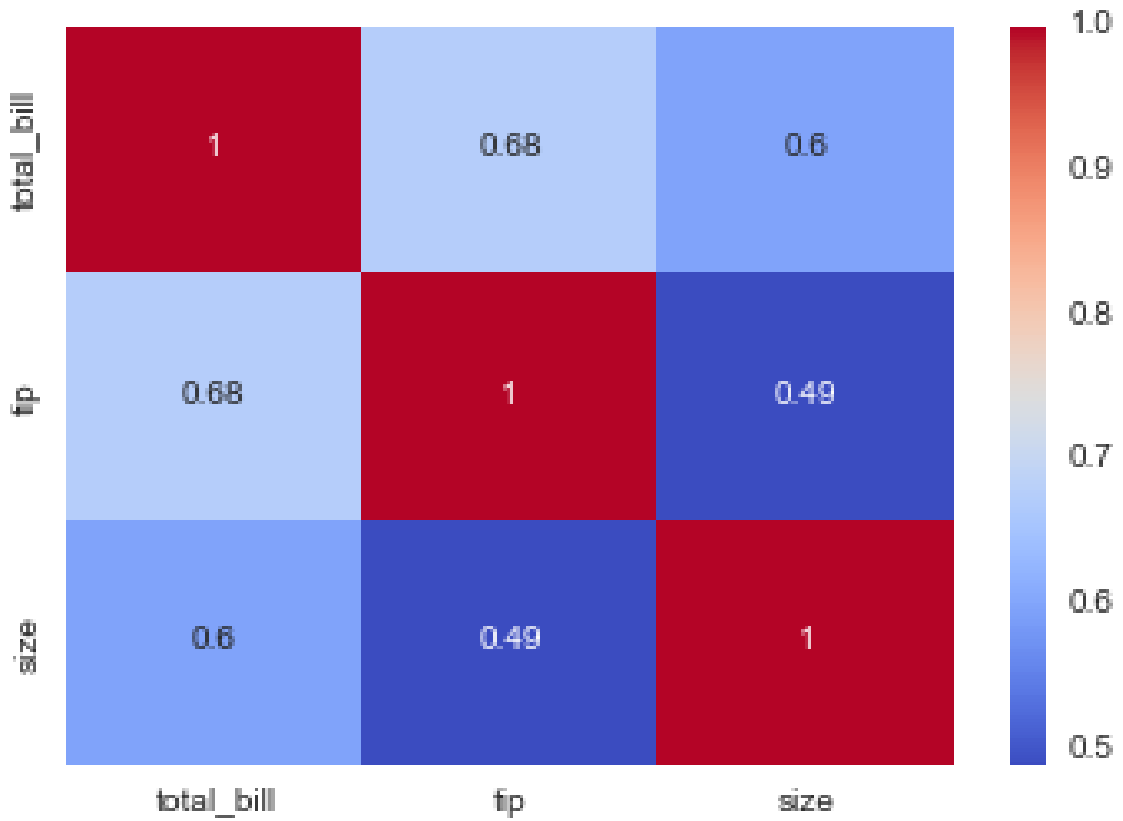
- The stripplot will draw a scatterplot where one variable is categorical. A strip plot can be drawn on its own, but it is also a good complement to a box or violin plot in cases where you want to show all observations along with some representation of the underlying distribution.



Matrix Plots

Matrix plots allow you to plot data as color-encoded matrices.

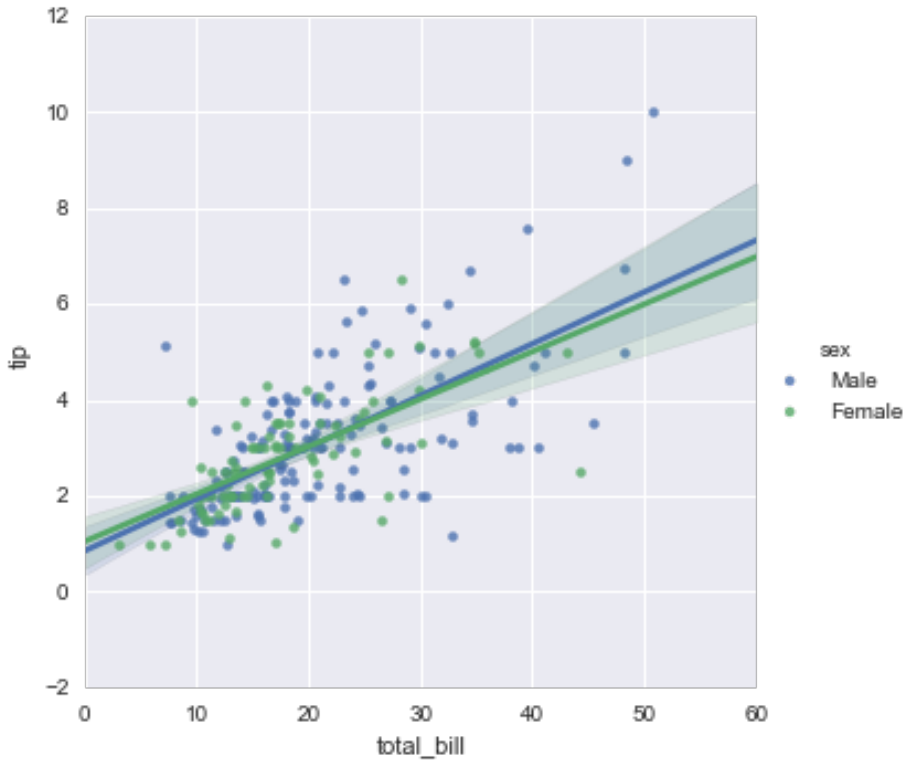
- Heatmap



Regression Plots

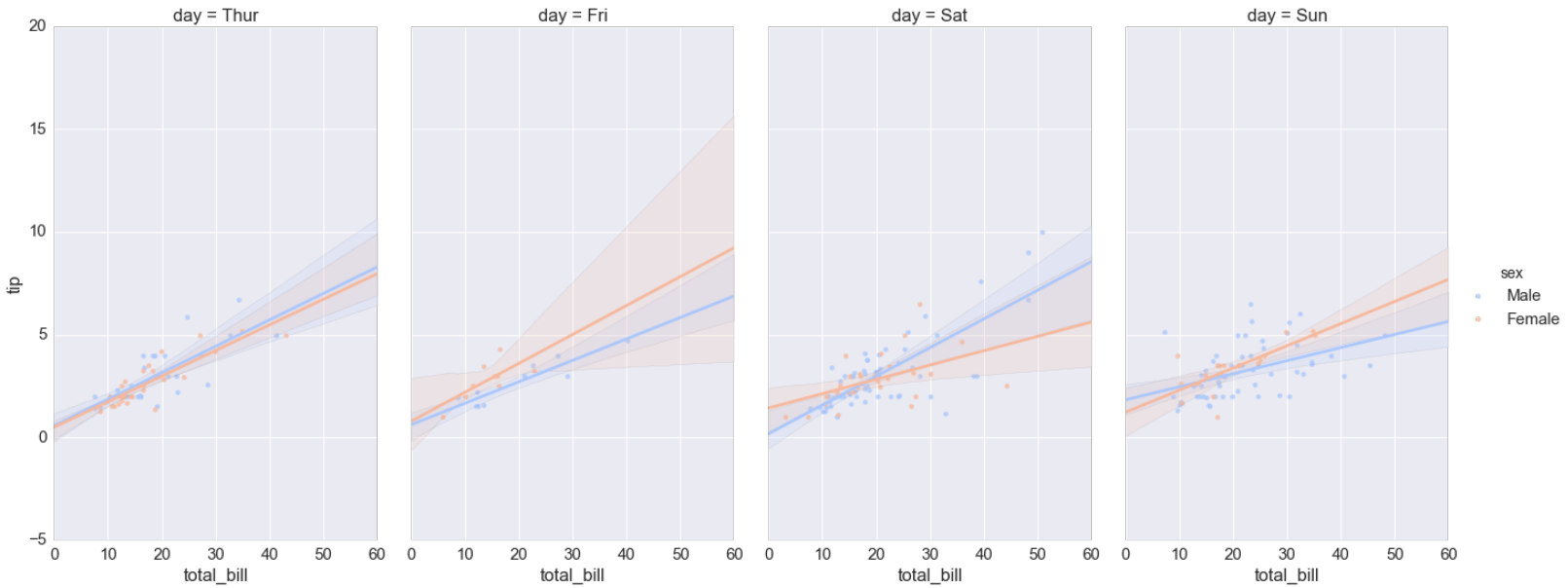
Seaborn has many built-in capabilities for regression plots, however we won't really discuss regression until the machine learning section of the course, so we will only cover the lmplot() function for now. lmplot allows you to display linear models.

- Box Plots



Aspect and Size

- Seaborn figures can have their size and aspect ratio adjusted with the size and aspect parameters:



Great Job
Thank yo