GraphLab Cheatsheet

Contents

[GraphLab Cheatsheet 1](#_Toc431436999)

[Environment 1](#_Toc431437000)

[Plot into the iPython notebook 1](#_Toc431437001)

[Images 1](#_Toc431437002)

[Show an image 1](#_Toc431437003)

[Plotting 2](#_Toc431437004)

[Box Whisker Plot 2](#_Toc431437005)

[Scatter Plot 2](#_Toc431437006)

[Regression 3](#_Toc431437007)

[Create a regression model 3](#_Toc431437008)

[Coefficients of a regression model 3](#_Toc431437009)

[SFrames 3](#_Toc431437010)

[Add a column to an Sframe 3](#_Toc431437011)

[Apply a function to an Sframe column 3](#_Toc431437012)

[Max of a column 3](#_Toc431437013)

[Mean of a column 4](#_Toc431437014)

[Random partition of an Sframe 4](#_Toc431437015)

[Summary of a column 4](#_Toc431437016)

[Visualize columns 5](#_Toc431437017)

# Environment

## Plot into the iPython notebook

graphlab.canvas.set\_target('ipynb')

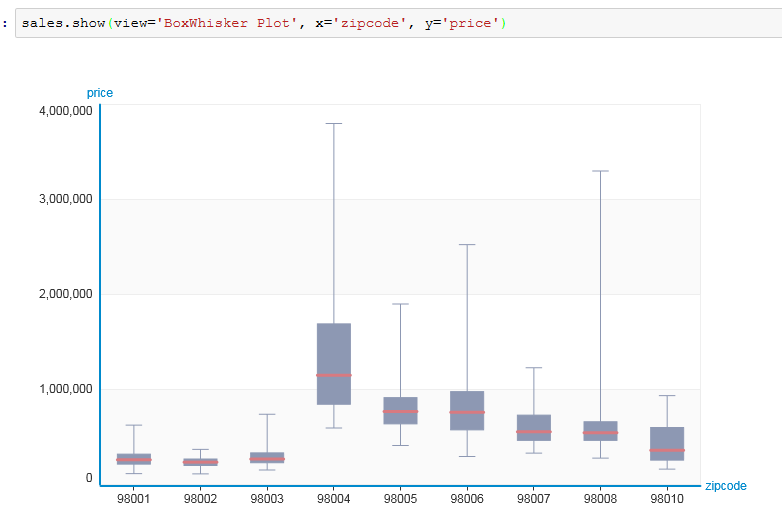
# Images

## Show an image

<img src="house-1925069082.jpg">

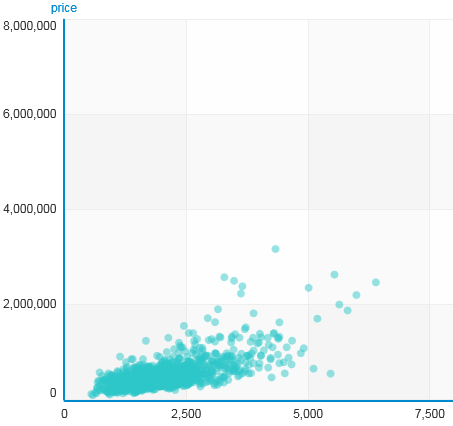
# Plotting

## Box Whisker Plot



## Scatter Plot

sales.show(view="Scatter Plot", x="sqft\_living", y="price")



# Regression

## Create a regression model

sqft\_model = graphlab.linear\_regression.create(train\_data, target='price', features=['sqft\_living'])

## Coefficients of a regression model

sqft\_model.get('coefficients')

# SFrames

## Add a column to an Sframe

sf['Full Name'] = sf['First Name'] + ' ' + sf['Last Name']

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **First Name** | **Last Name** | **Country** | **age** | **Full Name** |
| Bob | Smith | United States | 24 | Bob Smith |
| Alice | Williams | Canada | 23 | Alice Williams |
| Malcolm | Jone | England | 22 | Malcolm Jone |

## Apply a function to an Sframe column

def transform\_country(country):

if country == 'USA':

return 'United States'

else:

return country

sf['Country'] = sf['Country'].apply(transform\_country)

## Max of a column

sf['age'].max()

## Mean of a column

sf['age'].mean()

## Random partition of an Sframe

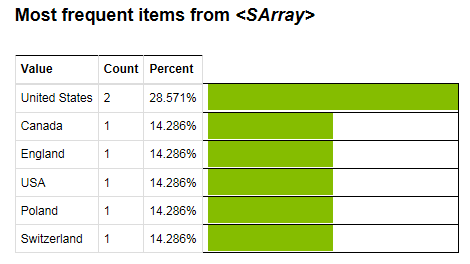
train\_data,test\_data = sales.random\_split(.8,seed=0)

## Row Count

sales.num\_rows()

## Summary of a column

sf['Country'].show()



## Visualize columns

