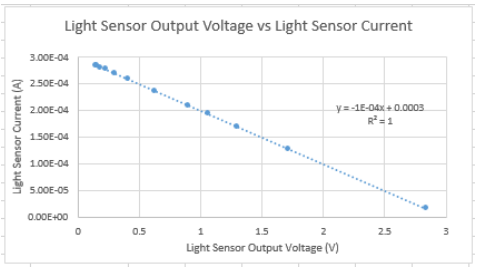
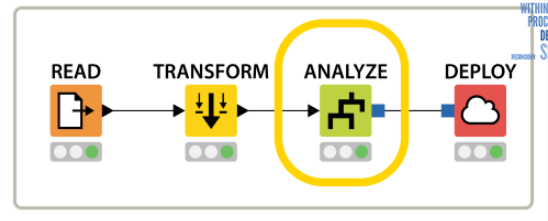
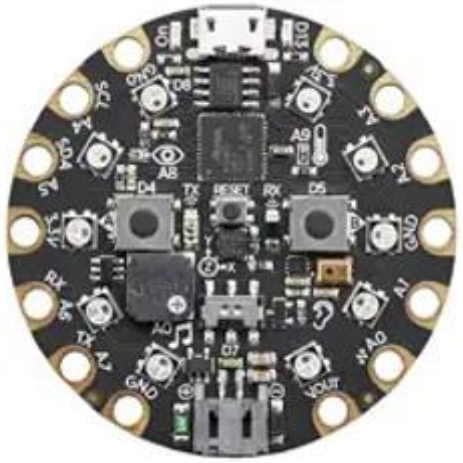
Prototyping Predictive Analytics Techniques

Class 2: Creating a Predictive Analytic Model







March 19, 2019

Don Wilcher

Presented by:

DesignNews

CEC CONTINUING EDUCATION CENTER

Digi-Key ELECTRONICS

Class 2: Creating a Predictive Analytic Model

Agenda

- The Predictive Analytics Process
- What is Linear Regression
- Exploring Orange
- Lab Project: Predicting the electrical behavior of a light sensor circuit.

The Predictive Analytics Process

There are seven steps to the Predictive Analytics Process.

1. **Define Project**
2. **Data Collection**
3. **Data Analysis**
4. **Statistics**
5. **Modeling**
6. **Deployment**
7. **Model Monitoring**

The Predictive Analytics Process…

1. **Project Definition** – Identifying the project outcomes based on the design objectives specified.
2. **Data Collection** – Acquiring and preparing data for the purpose of predicting future trends for processing and monitoring systems.
3. **Data Analysis** – The process of reviewing data results using statistical methods.

The Predictive Analytics Process…

4. Statistics – The validation of assumptions and hypothesis using statistical methods.

5. Modeling – The ability to create predictive models that can accurately forecast future trends.

6. Deployment – To send out the predictive model for use in decision making processes, tasks and activities.

7. Model Monitoring – The management and monitoring the predictive performance of the analytical model.

The Predictive Analytics Process…

Predictive Analytics (PA) Process Model

What is Linear Regression?

Linear Regression – Identifying the project outcomes based on the design objectives specified.

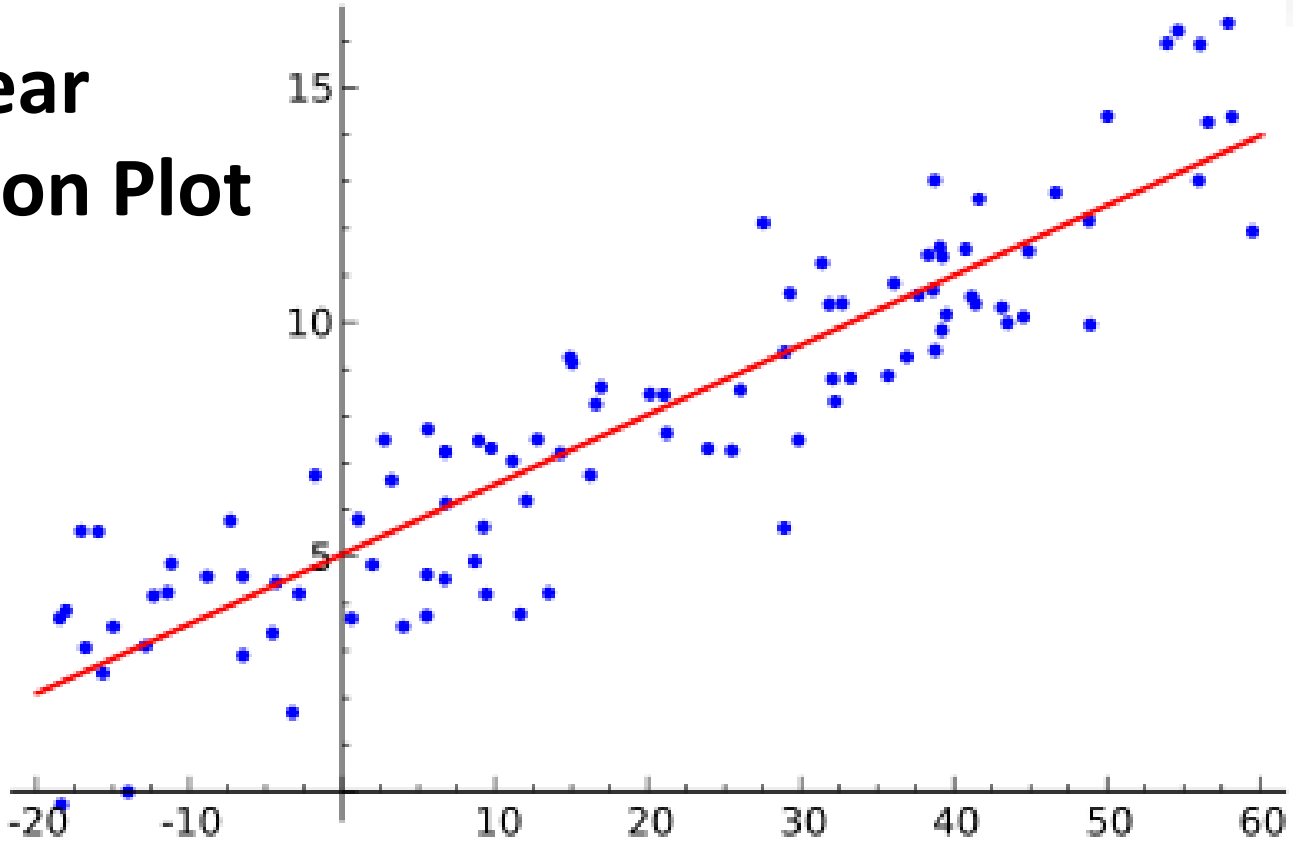
FAQs:

- The focus is on predicting a value of Y having been given the value of X.
- The regression line is called the *prediction line*.
- The regression line is a straight line that lies closet to all the points in the scatter plot.

Source:
Sprinthall, R. (1987). Basic statistical analysis, 3rd ed. Englewood Cliffs, NJ: Prentice Hall Inc.

What is Linear Regression?. . .

Linear Regression Plot



Question 1

True or False: There are eight steps for the Predictive Analytics Model process.

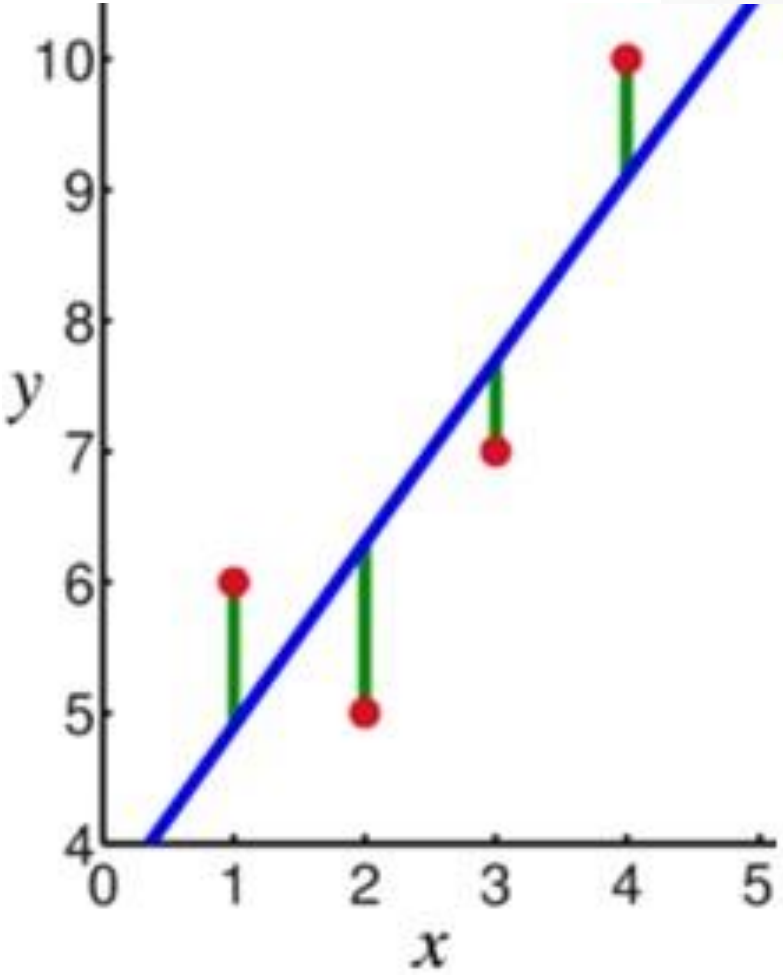
What is Linear Regression?. . .

Linear Regression Plot Anatomy

a)Red dots: observation

b)Green line: deviation

c)Blue line: *y* vs *x*



What is Linear Regression?. . .

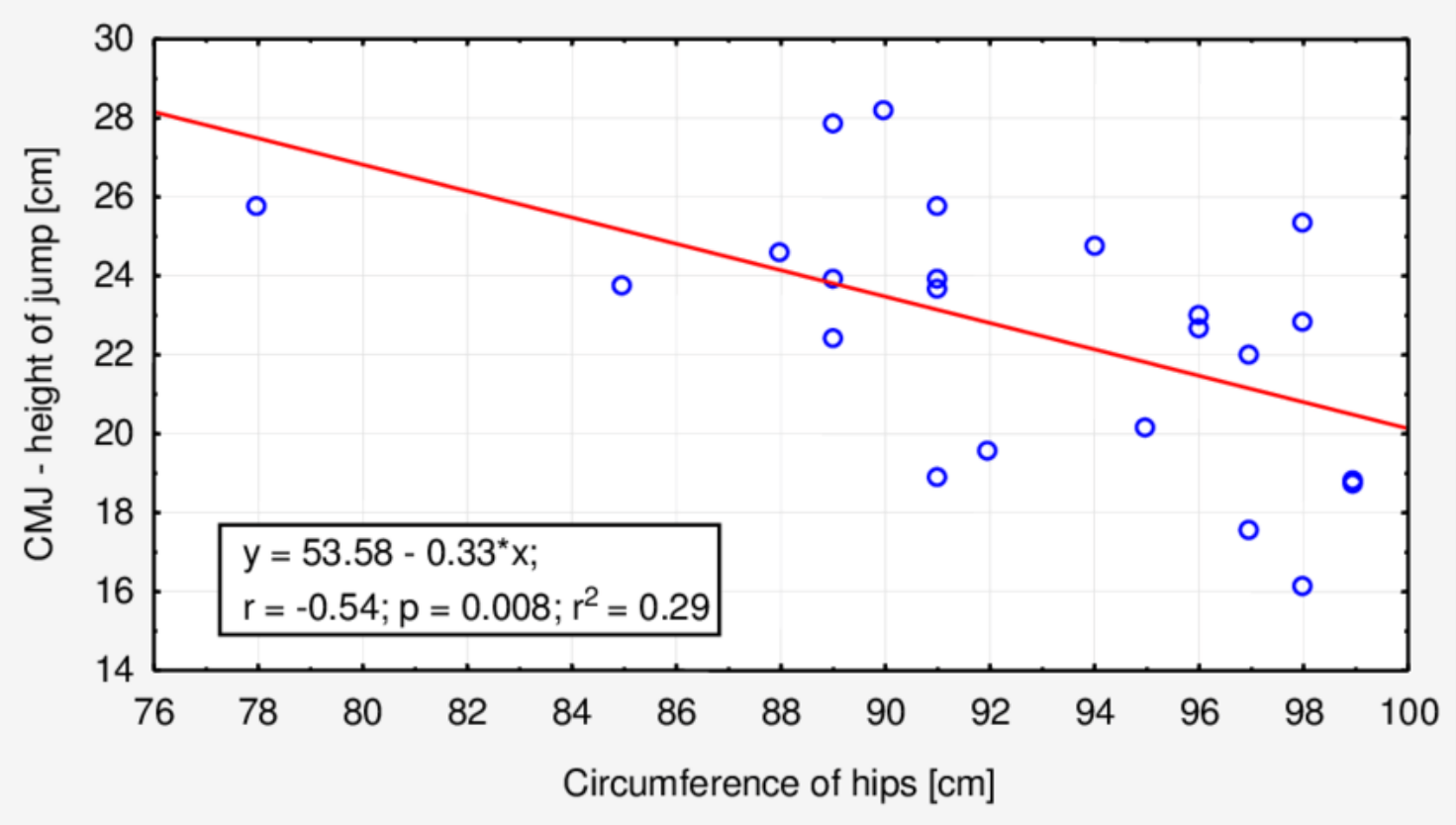
The Extent of the Scatter Around the Regression

- The closer the points on the scatter plot cluster around the regression line, the higher is the resulting correlation between x and y.
- The closer the points to the regression line, the more accurate is the resulting prediction.
- The higher the correlation, the closer the scatter points cluster around the regression line.

Source:

Sprinthall, R. (1987). *Basic statistical analysis*, 3rd ed. Englewood Cliffs, NJ: Prentice Hall Inc.

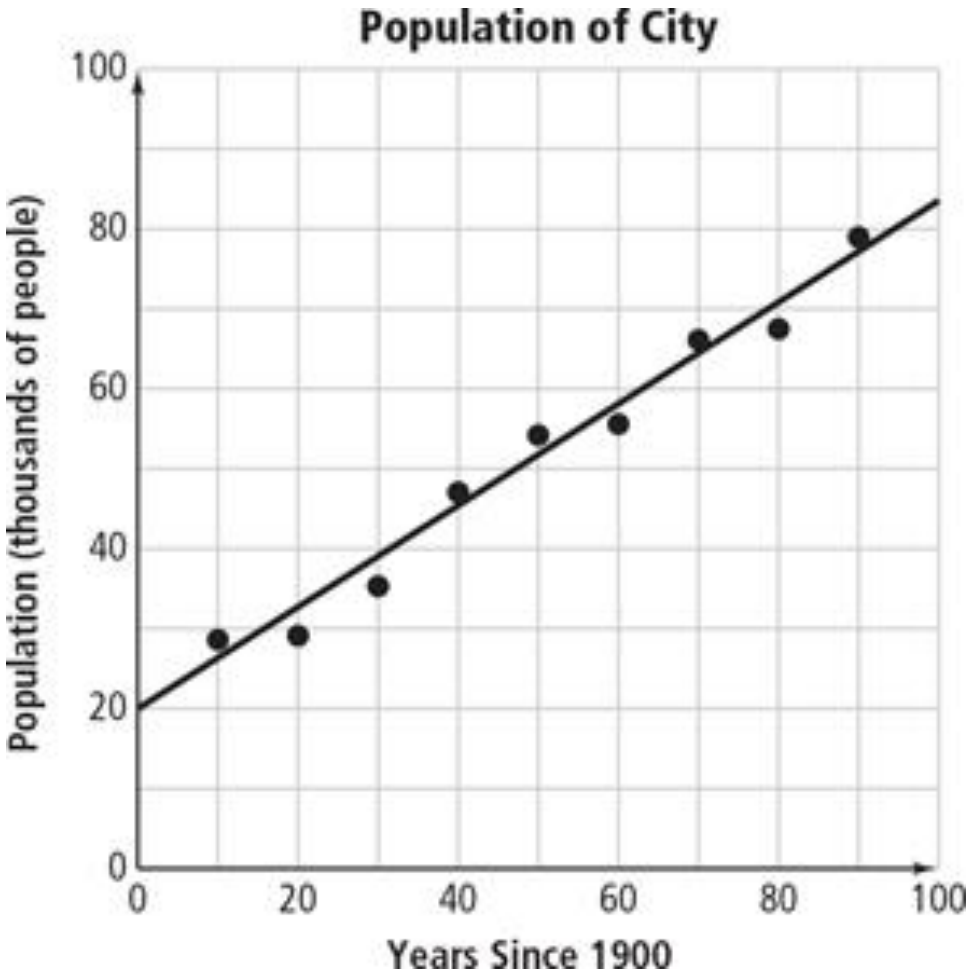
What is Linear Regression?. . .



Source:

https://www.researchgate.net/publication/309526456_Application_of_the_index_of_balance-stiffness_for_evaluation_of_the_process_of_maintaining_body_balance/figures?lo=1

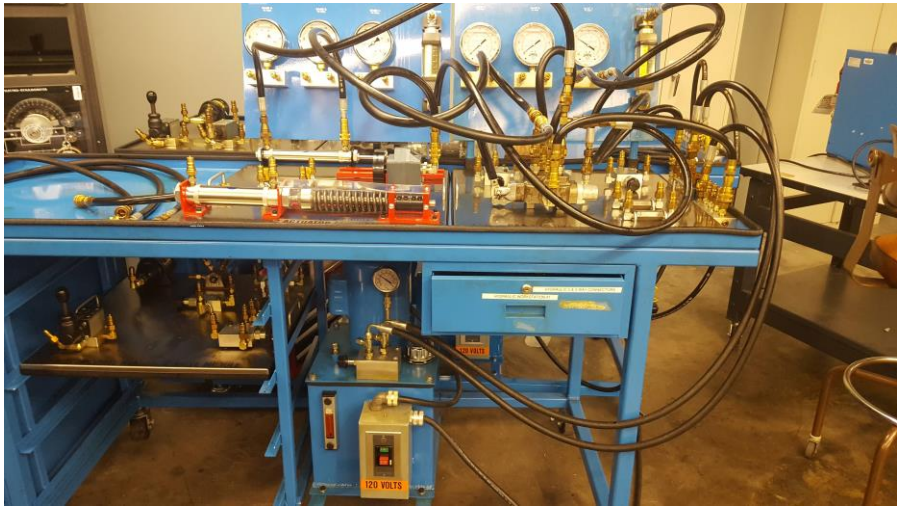
What is Linear Regression?. . .

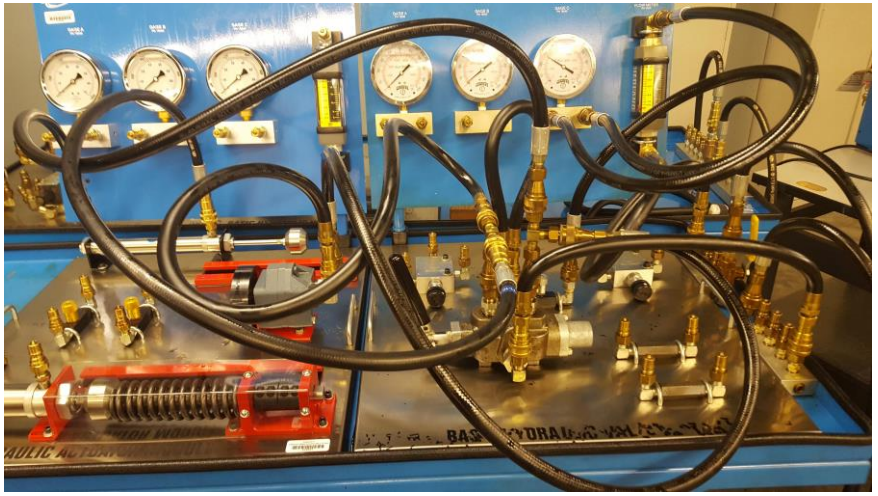


http://flipbooks.pearsonschool.com/texasreview/mathematics/digits/TX_Digits_HomeworkHelper_HTML_Files/Grade%208/Volume%202/page_385.html

What is Linear Regression?. . .

Predictive Maintenance: Hydraulic System

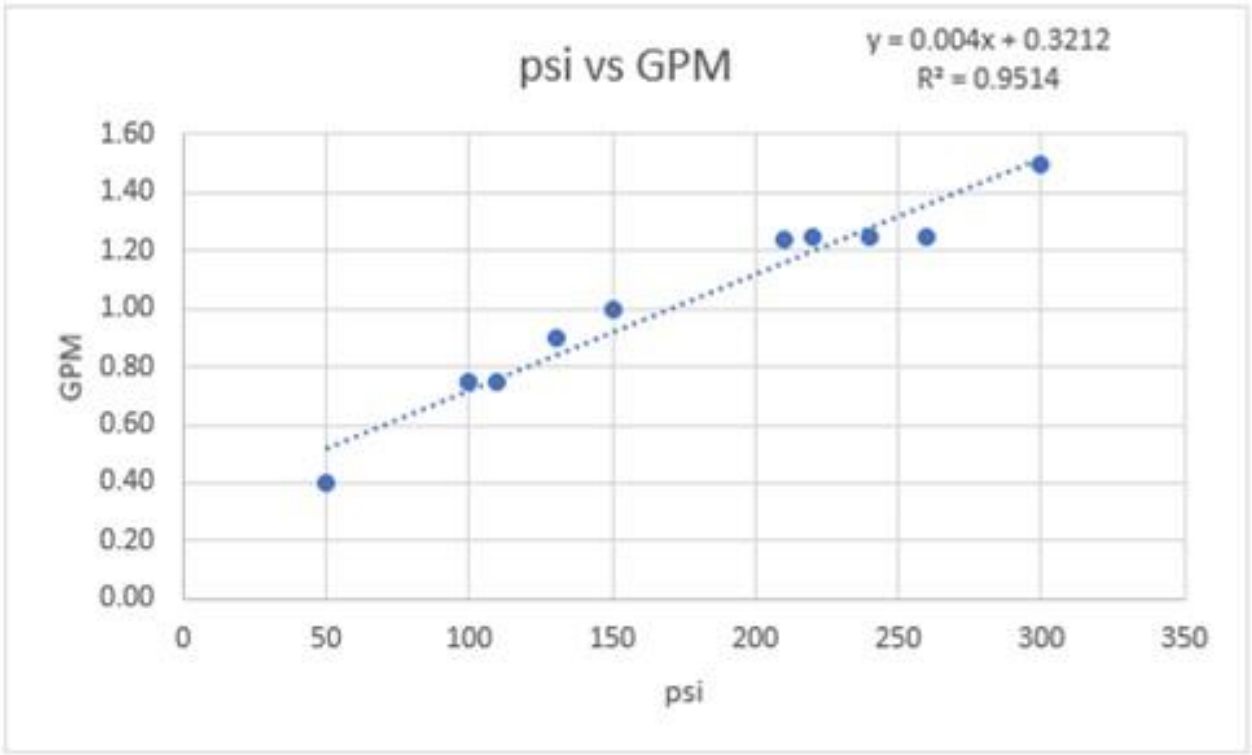




What is Linear Regression?. . .

Predictive Maintenance: Hydraulic System Data

psi	GPM
300	1.50
260	1.25
240	1.25
220	1.25
210	1.24
150	1.00
130	0.90
110	0.75
100	0.75
50	0.40



What is Linear Regression?. . .

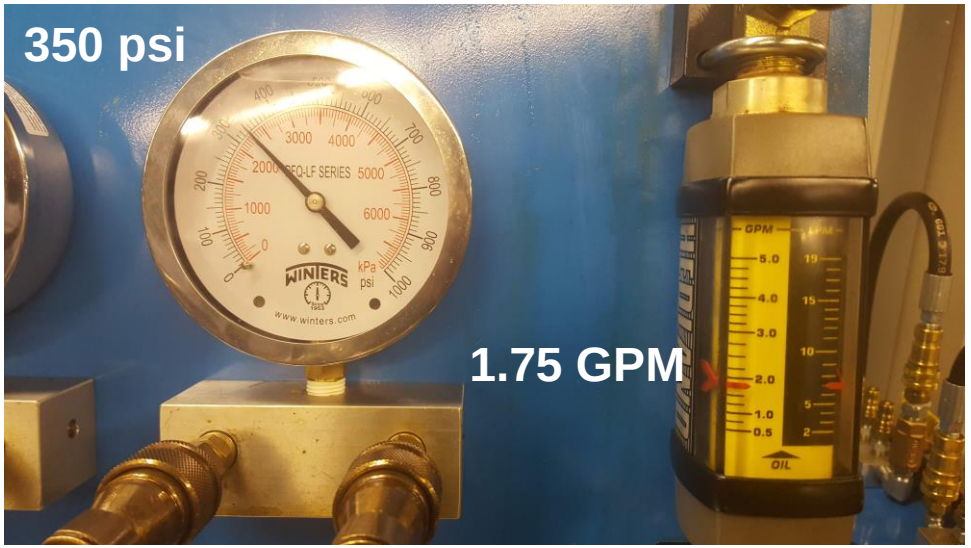
Predictive Maintenance: Hydraulic System Prediction

What would be the fluid flow rate if the system pressure is 350 psi?
Solution:

x is the input variable
y is the outcome variable

$$y = 0.004x + 0.3212$$
$$y = 0.004\ (350psi) + 0.3212$$
$$y = 1.72\ GPM\ (Predictive\ Value)$$

System Measurement



Question 2

True or False: Linear Regression is a complex Machine Learning Model for Predictive Analytics.

What is Linear Regression?. . .

Linear Regression based on this simple equation.

$$y = mx + b$$

output — slope — input — y-intercept

Note:

and are coefficients for the linear equation

Source:
https://www.sas.com/en_us/insights/analytics/predictive-analytics.html

Exploring Orange

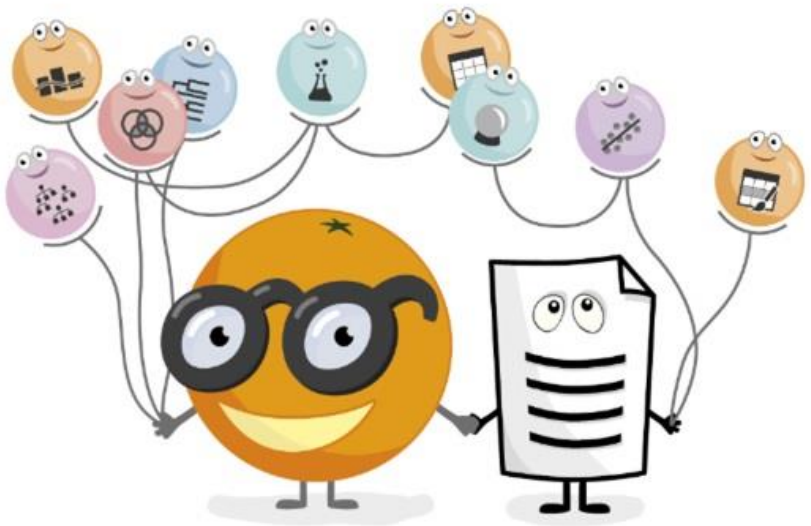
orange

Features Screenshots Download Docs Blog Training Donate

Data Mining
Fruitful and Fun

Open source machine learning and data visualization for novice and expert. Interactive data analysis workflows with a large toolbox.

Download Orange

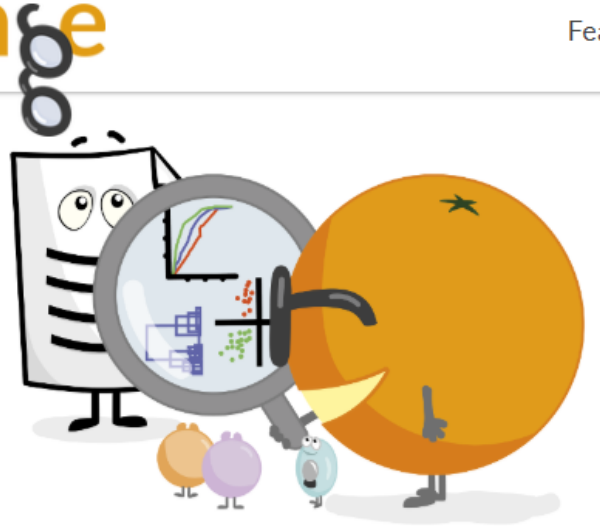


Source:
https://orange.biolab.si/

Exploring Orange. . .

orange

Features Screenshots Download Docs Blog Training



Interactive Data Visualization

Perform simple data analysis with clever data visualization. Explore statistical distributions, box plots and scatter plots, or dive deeper with decision trees, hierarchical clustering, heatmaps, MDS and linear projections. Even your multidimensional data can become sensible in 2D, especially with clever attribute ranking and selections.

Learn More

Visual Programming

Interactive data exploration for rapid qualitative analysis with clean visualizations. Graphic user interface allows you to focus on exploratory data analysis instead of coding, while clever defaults make fast prototyping of a data analysis workflow extremely easy. Place widgets on the canvas, connect them, load your datasets and harvest the insight!

Learn More Watch Video

Source:
https://orange.biolab.si/

Exploring Orange. . .

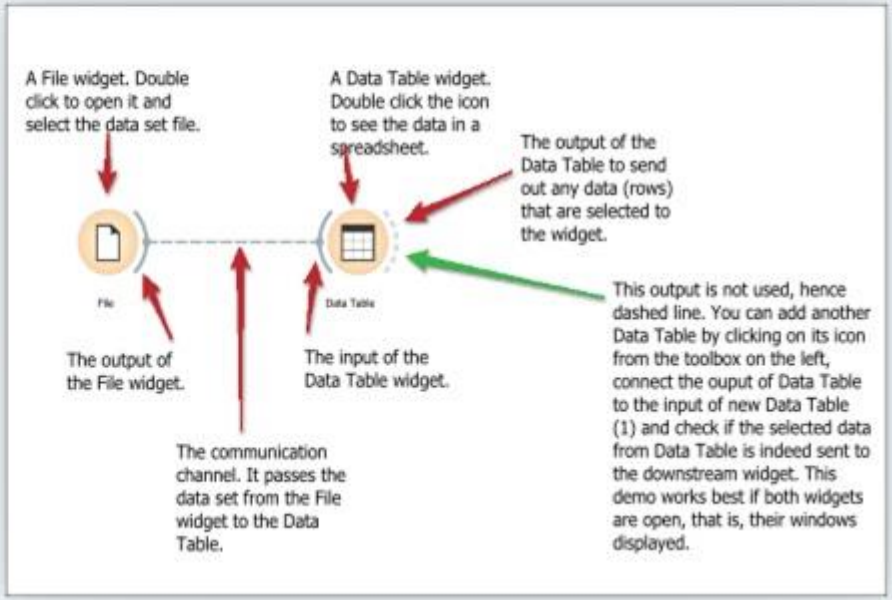
Example Workflows

Example Workflows

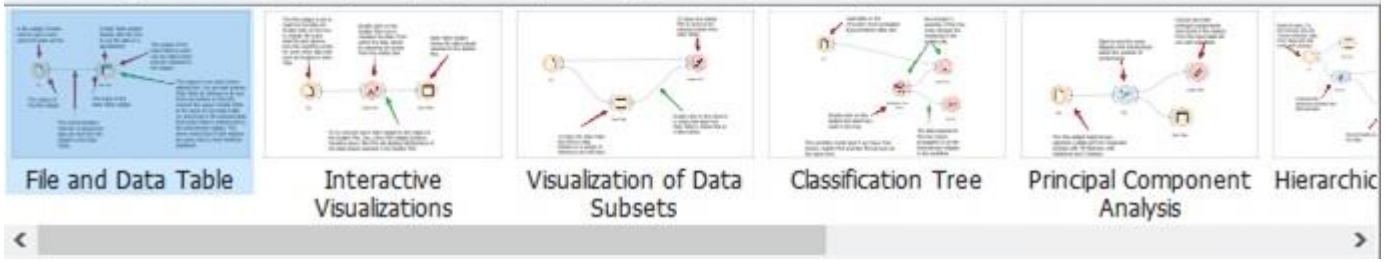
File and Data Table

The basic data mining units in Orange are called widgets. There are widgets for reading the data, preprocessing, visualization, clustering, classification and others. Widgets communicate through channels. Data mining workflow is thus a collection of widgets and communication channels.

In this workflow, there is a File widget that reads the data. File widget communicates this data to Data Table widget that shows the data spreadsheet. Notice how the output of the file widget is connected to the input of the Data Table widget. In Orange, the outputs of the widgets are on the right, and the inputs on the left of the widget.



Path: C:\Anaconda3\lib\site-packages\Orange\canvas\application\workflows\110-file-and-data-table-widget.ows



Exploring Orange. . .

orange

Home Screenshots Download Docs Blog Training

Documentation

Visual Programming

Getting started
YouTube tutorials
Loading your data
Widget catalog

Development

Widget development
Example addon

Python Library

Tutorial
Reference
Orange 2.7 documentation

Source:
https://orange.biolab.si/docs/

Exploring Orange. . .

Cricket's Chirp Rate vs Temperature Data: CSV file

```
Chirp Rate,Temperature
20,88.6
16,71.6
19.8,93.3
18.4,84.3
17.1,80.6
15.5,75.2
14.7,69.7
17.1,82
15.4,69.4
16.3,83.3
15,79.6
17.2,82.6
16,80.6
17,83.5
14.4,76.3
```

Question 3

Orange is an open source machine learning and _____ visualization tool for novice and expert developers.

Exploring Orange. . .

orange

Home Screenshots Download Docs Blog Training

Quick Search

Data

File Datasets SQL Table Data Table Paint Data Data Info Data Sampler

Select Columns Select Rows Rank Merge Data Concatenate Transpose Randomize

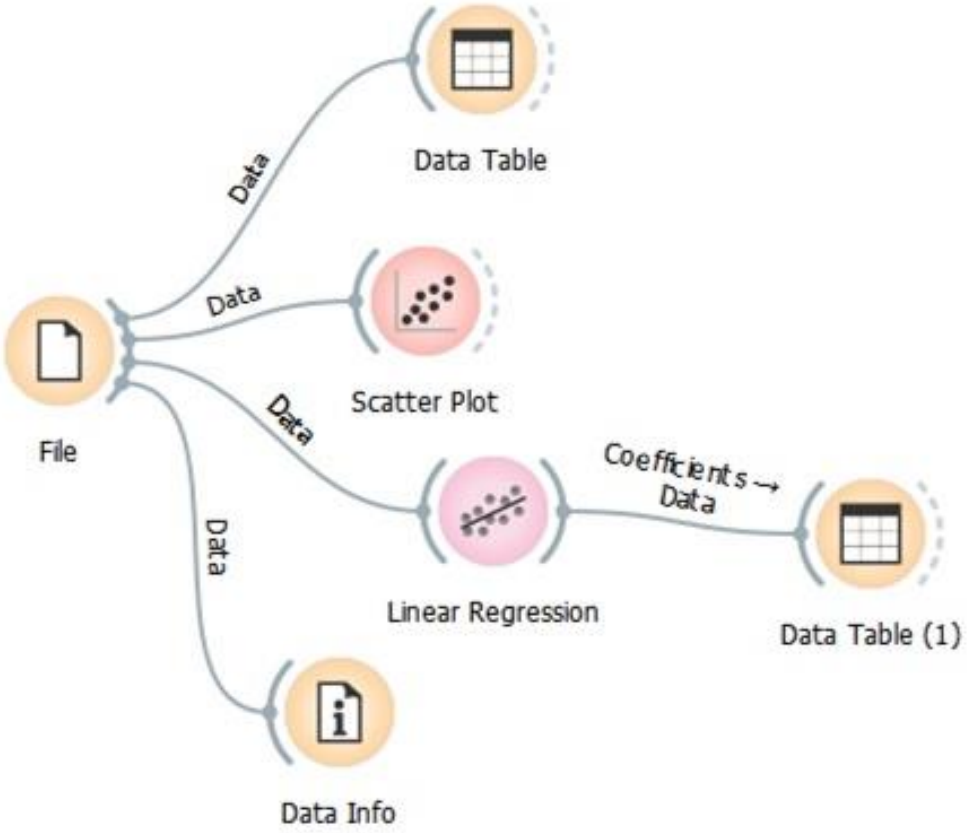
Preprocess Impute Outliers Edit Domain Python Script Color Continuize

Create Class Discretize Feature Constructor Purge Domain Save Data

Source:
https://orange.biolab.si/toolbox/

Exploring Orange. . .

Analyzing a Cricket's Chirp Rate vs Temperature: Workflow



Exploring Orange. . .

Cricket's Chirp Rate vs Temperature Data: CSV file

```
Chirp Rate,Temperature
20,88.6
16,71.6
19.8,93.3
18.4,84.3
17.1,80.6
15.5,75.2
14.7,69.7
17.1,82
15.4,69.4
16.3,83.3
15,79.6
17.2,82.6
16,80.6
17,83.5
14.4,76.3
```

Question 4

Using slide 23: In configuring the input file, what feature is used as the target for the machine learning model in Orange?

Exploring Orange. . .

Data Table

Info

15 instances (no missing values)
1 feature (no missing values)
Continuous target variable (no missing values)
No meta attributes

Variables

- [x] Show variable labels (if present)
- [ ] Visualize continuous values
- [x] Color by instance classes

Selection

- [x] Select full rows

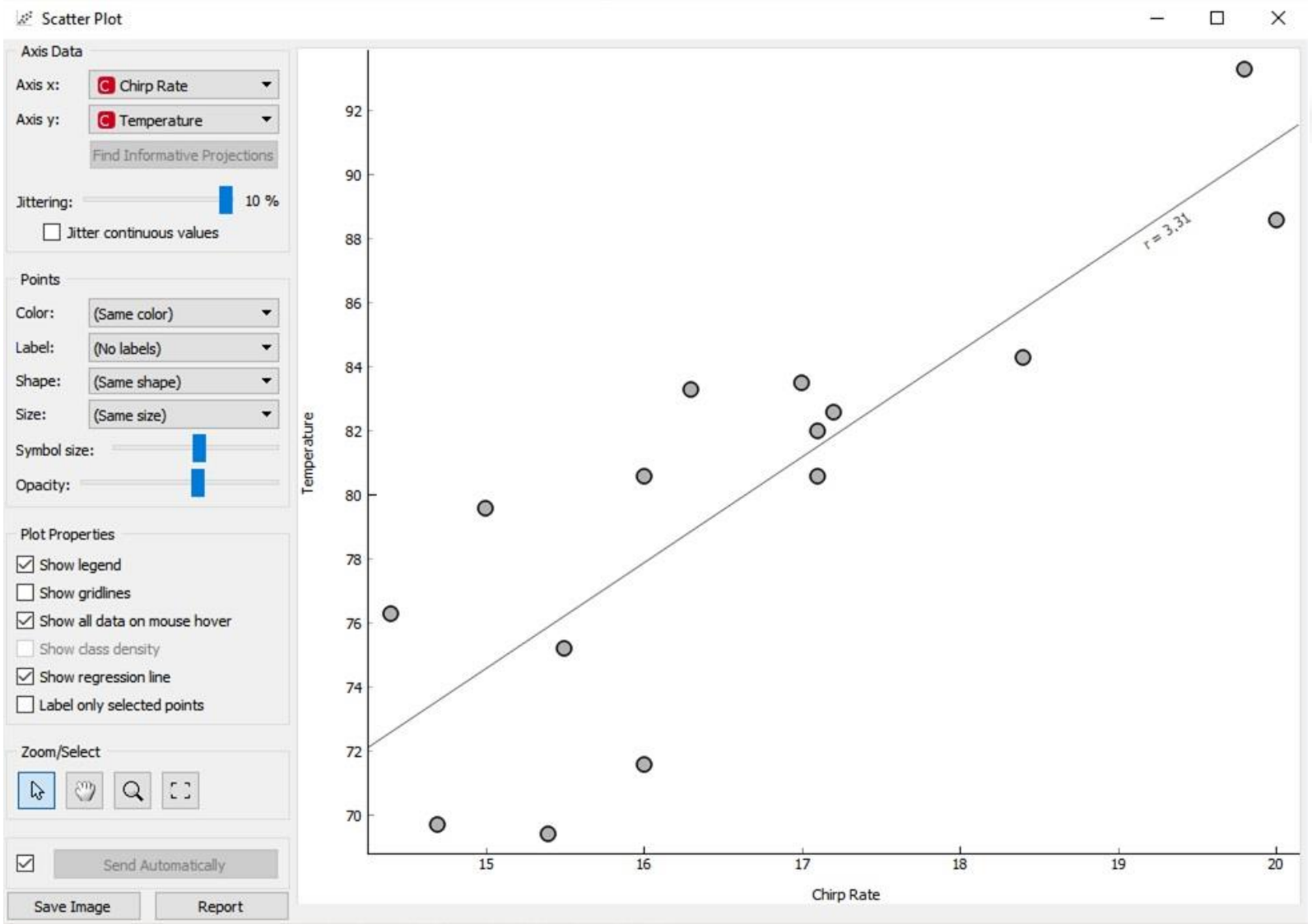
Restore Original Order

Report

- [x] Send Automatically

	Temperature	Chirp Rate
1	88.600	20.000
2	71.600	16.000
3	93.300	19.800
4	84.300	18.400
5	80.600	17.100
6	75.200	15.500
7	69.700	14.700
8	82.000	17.100
9	69.400	15.400
10	83.300	16.300
11	79.600	15.000
12	82.600	17.200
13	80.600	16.000
14	83.500	17.000
15	76.300	14.400

Exploring Orange. . .



Exploring Orange. . .

Data Table (1)

Info

2 instances (no missing values)
1 feature (no missing values)
No target variable.
1 meta attribute (no missing values)

Variables

- [x] Show variable labels (if present)
- [ ] Visualize continuous values
- [x] Color by instance classes

Selection

- [x] Select full rows

	name	coef
1	intercept	24.9660144
2	Chirp Rate	3.3057614

Restore Original Order

Report

Send Automatically

Linear Regression equation.

$$y = 3.3057614x + 24.9660144$$

Exploring Orange. . .

Predicting a new temperature with a future chirp rate example.

Given:

$x = 21Hz$

$$y = 3.3057614x + 24.9660144$$

$$y = 3.3057614(21) + 24.9660144$$

$$y = 94.3869744\ °F$$

Question 5

Using slide 31: The numbers in the linear regression expression are the _____ of the equation.

Predicting the Electrical Behavior of a Light Sensor

Lab Project Objectives:

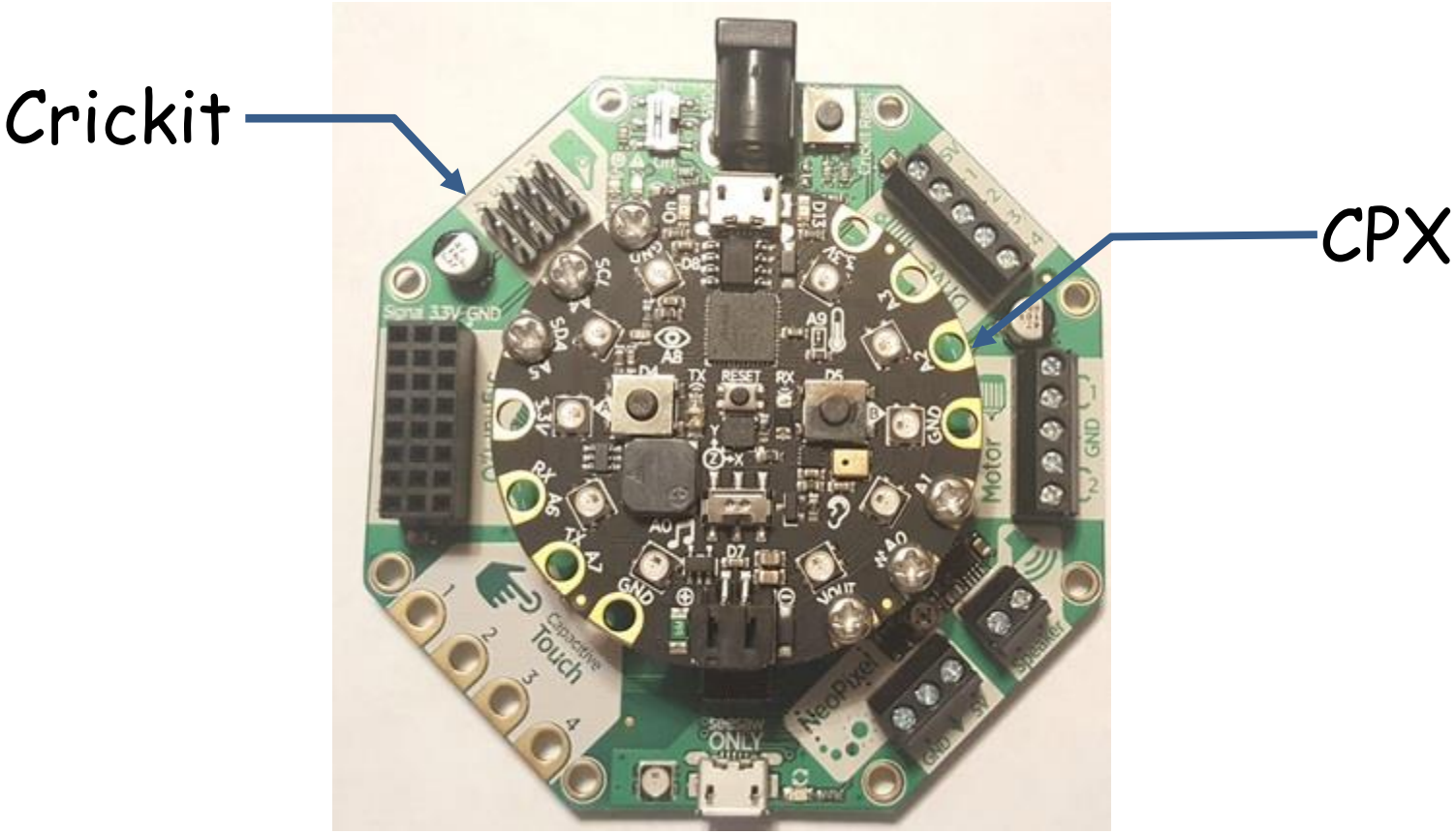
a. Learn how to collect data from the Circuit Playground Express (CPX) Light Sensor.

b. Learn how to create Common Separated Values file.

c. Learn how to build a Predictive Model using Orange.

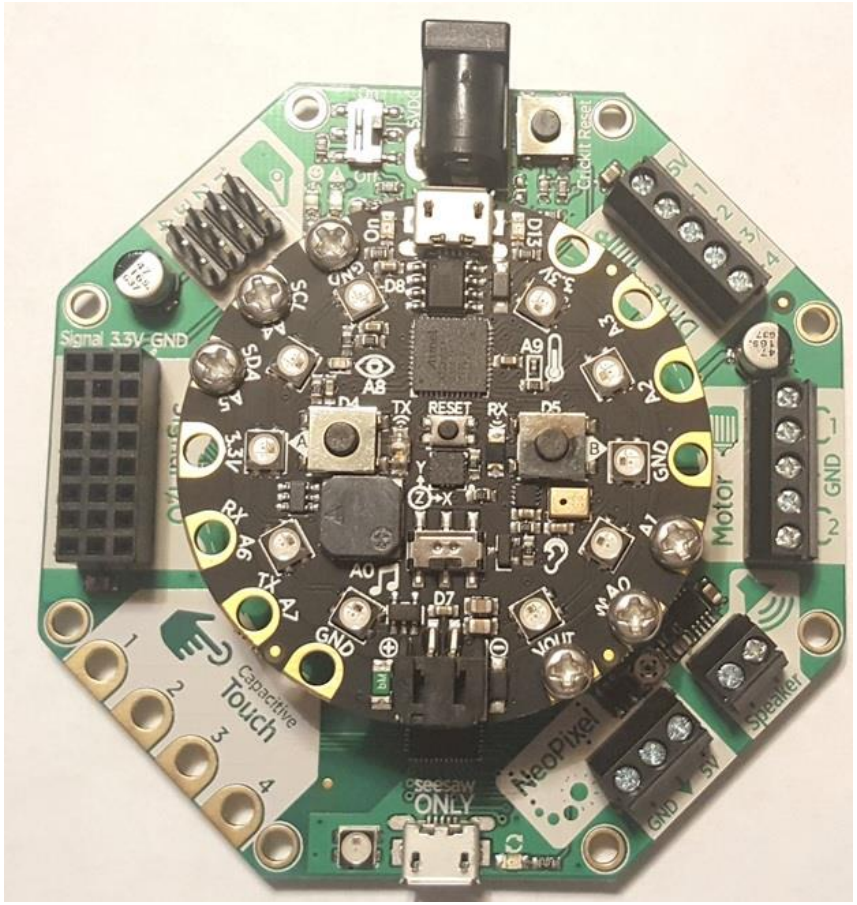
Predicting the Electrical Behavior of a Light Sensor Circuit. . .

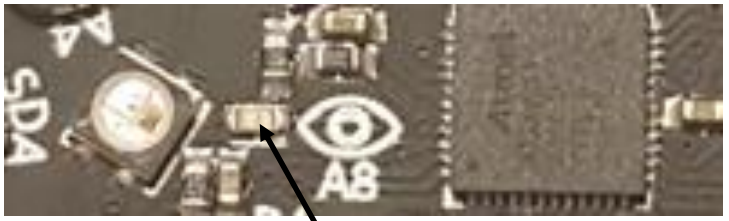
CPX mounted onto a Crickit.



Predicting the Electrical Behavior of a Light Sensor Circuit. . .

CPX mounted onto a Crickit.



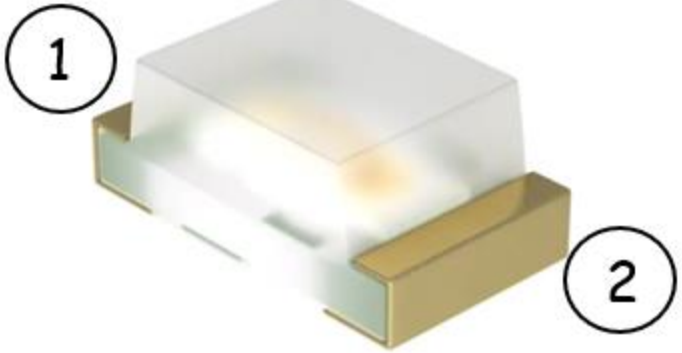


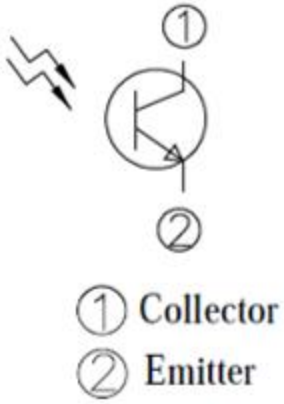
CPX
Light
Sensor

Predicting the Electrical Behavior of a Light Sensor Circuit. . .

CPX Light Sensor

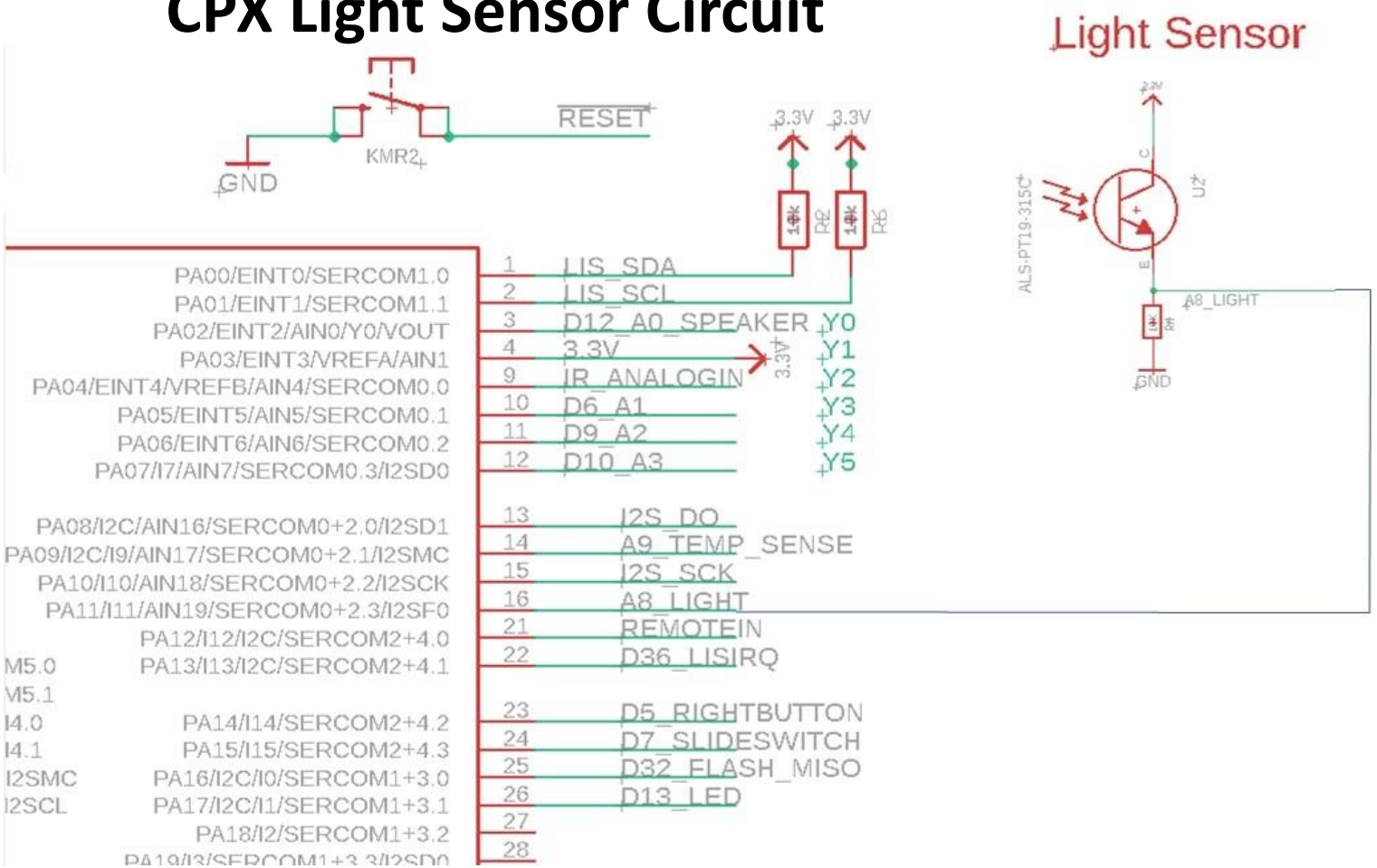
ALS-PT19-315C/L177/TR8
Surface Mount Light Sensor





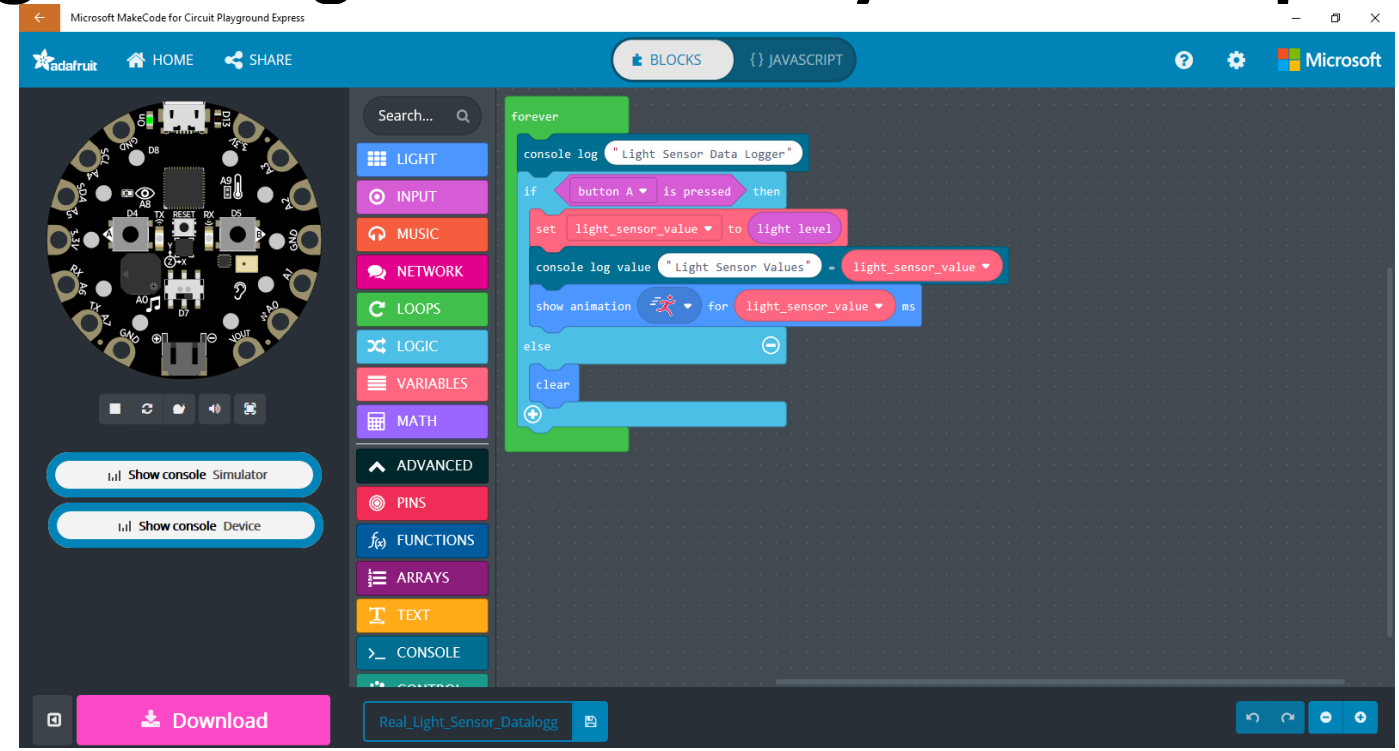
Predicting the Electrical Behavior of a Light Sensor Circuit. . .

CPX Light Sensor Circuit



Predicting the Electrical Behavior of a Light Sensor Circuit. . .

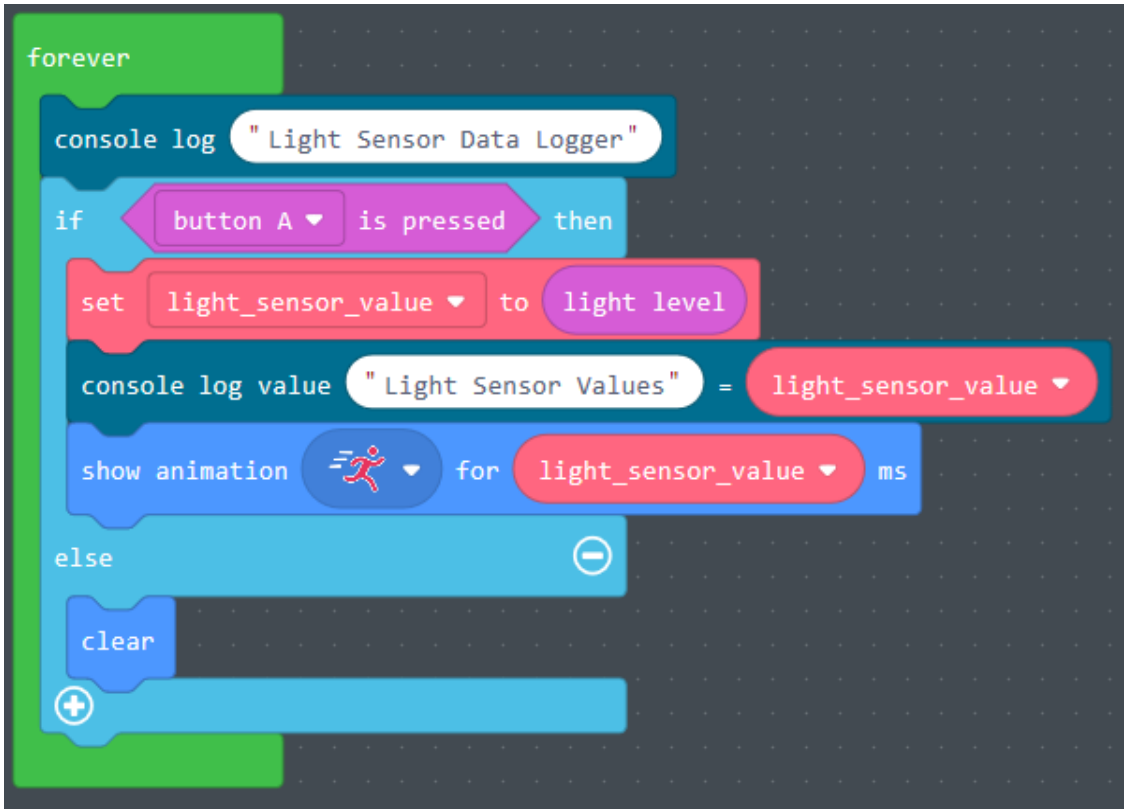
CPX Light Sensor Lab: Microsoft Makecode programming for Circuit Play Ground Express



Code Link:
https://www.tinkercad.com/things/aUlgFjfHDw5-electronic-light-sensor-circuit/editel

Predicting the Electrical Behavior of a Light Sensor Circuit. . .

CPX Light Sensor Lab: Microsoft Makecode Blockly Code



Predicting the Electrical Behavior of a Light Sensor Circuit. . .

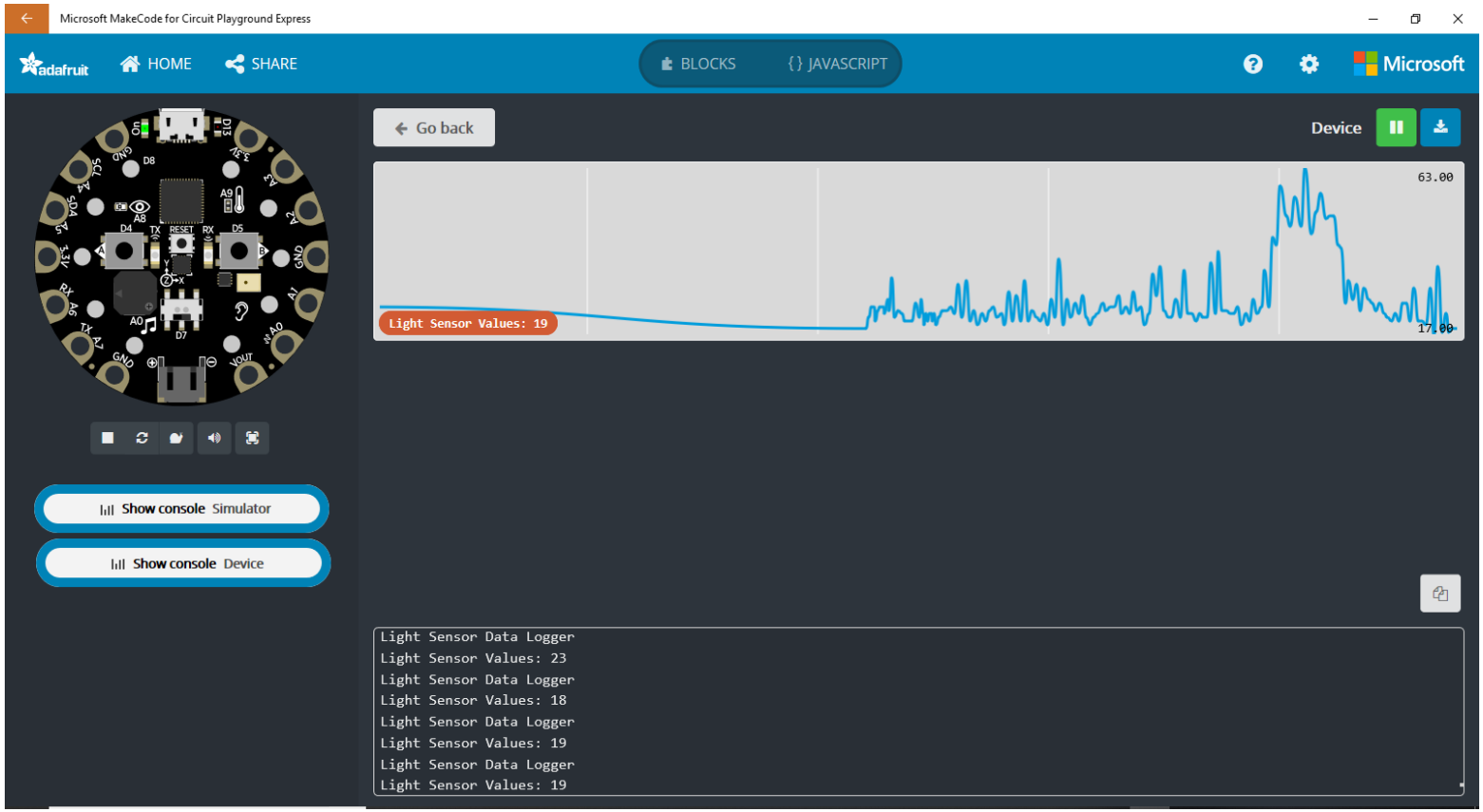
CPX Light Sensor Lab: Microsoft Makecode Javascript

```
let light_sensor_value = 0
forever(function () {
    console.log("Light Sensor Data Logger")
    if (input.buttonA.isPressed()) {
        light_sensor_value = input.lightLevel()
        console.logValue("Light Sensor Values", light_sensor_value)
        light.showAnimation(light.runningLightsAnimation, light_sensor_value)
    } else {
        light.clear()
    }
})

```

Predicting the Electrical Behavior of a Light Sensor Circuit. . .

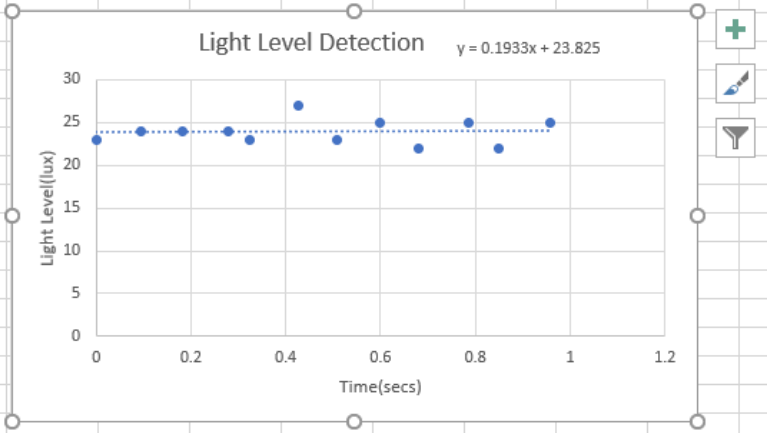
Running the application code on the CPX: Console Results



Predicting the Electrical Behavior of a Light Sensor Circuit. . .

Linear Regression Model of the CPX Light Sensor

	time (source1)	Light Sensor Values
2	0	23
3	0.093	24
4	0.183	24
5	0.278	24
6	0.324	23
7	0.426	27
8	0.509	23
9	0.6	25
10	0.681	22
11	0.787	25
12	0.849	22
13	0.957	25
14	1.021	24
15	1.122	24
16	1.186	24
17	1.266	23
18	1.369	27
19	1.422	22
20	1.531	26
21	1.603	23
22	1.707	30
23	1.784	24
24	4.712	22
25	4.795	23
26	4.864	35
27	4.946	24
28	5.057	29
29	5.105	22



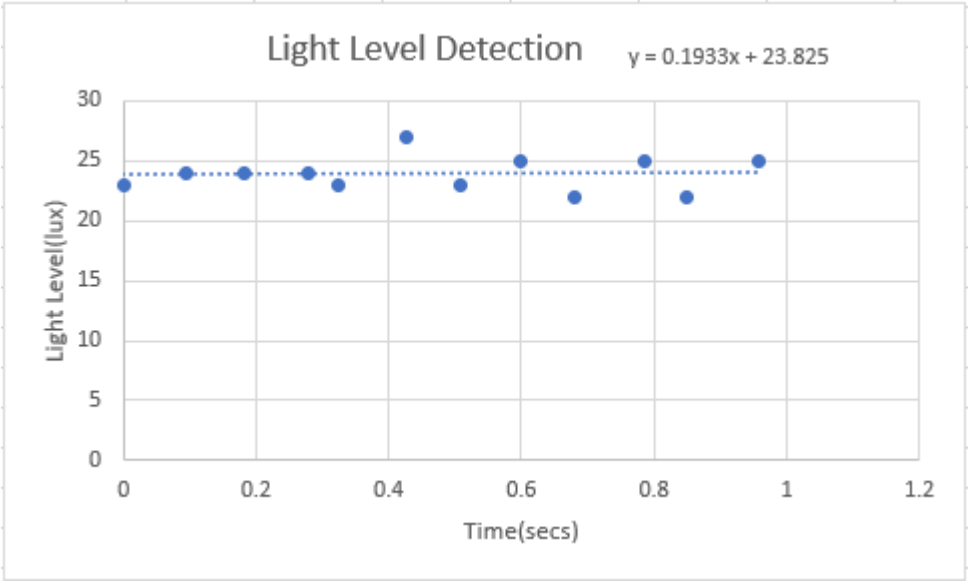
Predicting the Electrical Behavior of a Light Sensor Circuit. . .

What would be the light level if the used is 1.021 sec?

Solution:

x is the input variable
y is the outcome variable

$$y = 0.1933x + 23.825$$
$$y = 0.1933(1.021s) + 23.825$$
$$y = 24 \text{ lux (Predictive Value)}$$



Predicting the Electrical Behavior of a Light Sensor Circuit. . .

Linear Regression Model of the CPX Light Sensor:

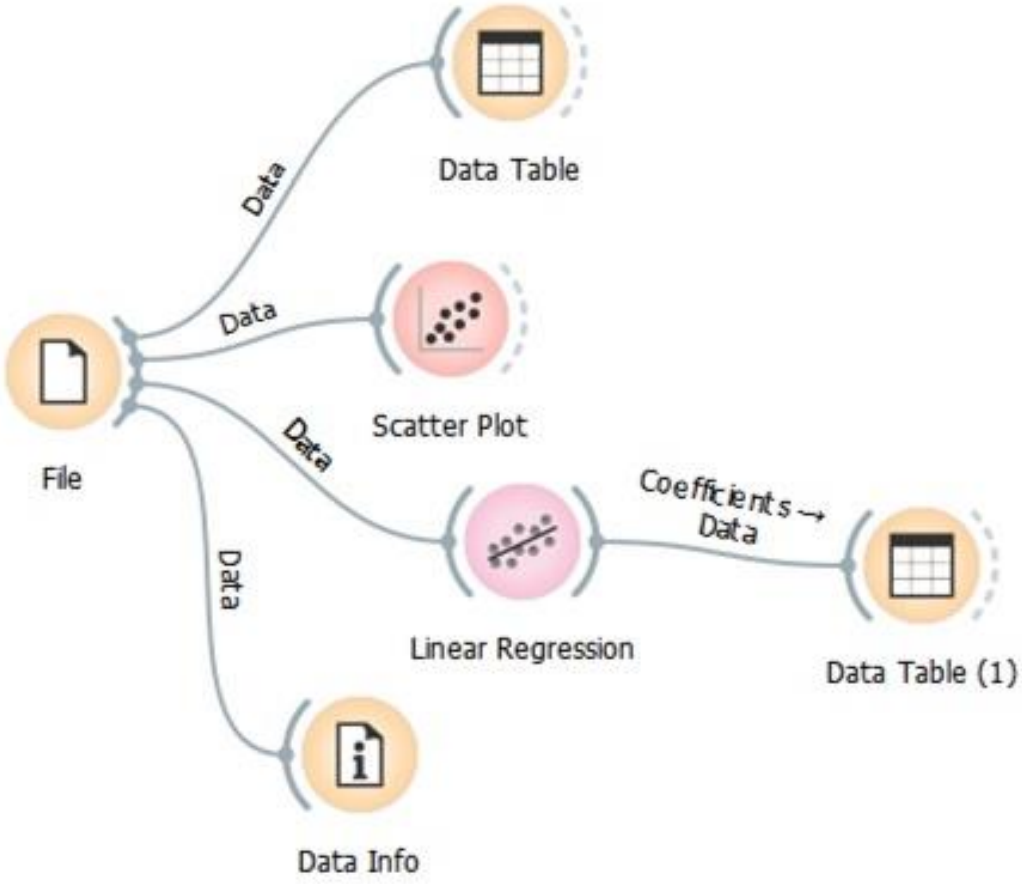
Measured Results

time (source1)	Light Sensor Values
0	23
0.093	24
0.183	24
0.278	24
0.324	23
0.426	27
0.509	23
0.6	25
0.681	22
0.787	25
0.849	22
0.957	25
1.021	24
1.122	24
1.186	24
1.266	23
1.369	27
1.422	22
1.531	26
1.603	23

Measured light level with CPX

Predicting the Electrical Behavior of a Light Sensor Circuit. . .

Light Sensor Circuit Predictive Model Workflow



Predicting the Electrical Behavior of a Light Sensor Circuit. . .

Light Sensor Circuit CSV File

```
time (sourcel),Light Sensor Values
0,23
0.093,24
0.183,24
0.278,24
0.324,23
0.426,27
0.509,23
0.6,25
0.681,22
0.787,25
0.849,22
0.957,25
1.021,24
1.122,24
```

Predicting the Electrical Behavior of a Light Sensor Circuit. . .

Light Sensor Circuit Data Table

Data Table

Info

14 instances (no missing values)
1 feature (no missing values)
Continuous target variable (no missing values)
No meta attributes

Variables

- [x] Show variable labels (if present)
- [ ] Visualize numeric values
- [x] Color by instance classes

Selection

- [x] Select full rows

Restore Original Order

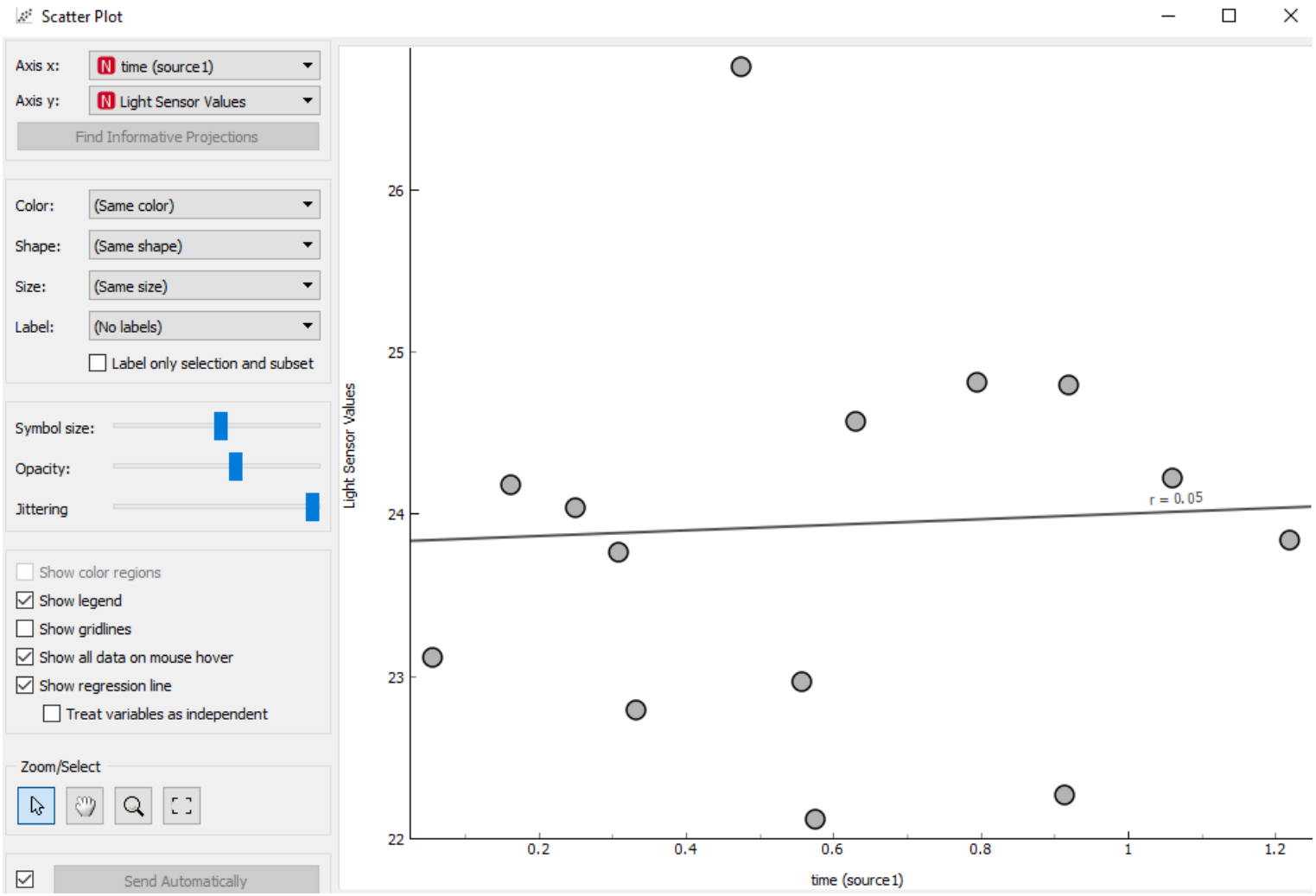
Send Automatically

	ight Sensor Value	time (source1)
1	23	0.000
2	24	0.093
3	24	0.183
4	24	0.278
5	23	0.324
6	27	0.426
7	23	0.509
8	25	0.600
9	22	0.681
10	25	0.787
11	22	0.849
12	25	0.957
13	24	1.021
14	24	1.122

Question 6

Using slide 48: In configuring the input file, what feature is used as the target for the machine learning model in Orange?

Predicting the Electrical Behavior of a Light Sensor Circuit. . .



Presented by:

Predicting the Electrical Behavior of a Light Sensor Circuit. . .

Data Table (1)

Info

2 instances (no missing values)
1 feature (no missing values)
No target variable.
1 meta attribute (no missing values)

Variables

- [x] Show variable labels (if present)
- [ ] Visualize numeric values
- [x] Color by instance classes

Selection

- [x] Select full rows

Restore Original Order

[x] Send Automatically

	name	coef
1	intercept	23.8318418
2	time (source1)	0.1729521

Linear Regression equation.

$$y=0.1729x + 23.8318$$

Predicting the Electrical Behavior of a Light Sensor Circuit. . .

Predicting a new light level value with the CPX light sensor circuit example.

Given:

$x = 1.186s$

$$y = 0.1729x + 23.8318$$

$$y = 0.1729(1.186s) + 23.8318$$

$$y = 24.03 \text{ lux}$$

time (source1)	Light Sensor Values
0	23
0.093	24
0.183	24
0.278	24
0.324	23
0.426	27
0.509	23
0.6	25
0.681	22
0.787	25
0.849	22
0.957	25
1.021	24
1.122	24
1.186	24
1.266	23
1.369	27
1.422	22
1.531	26
1.603	23

Measured light level with CPX