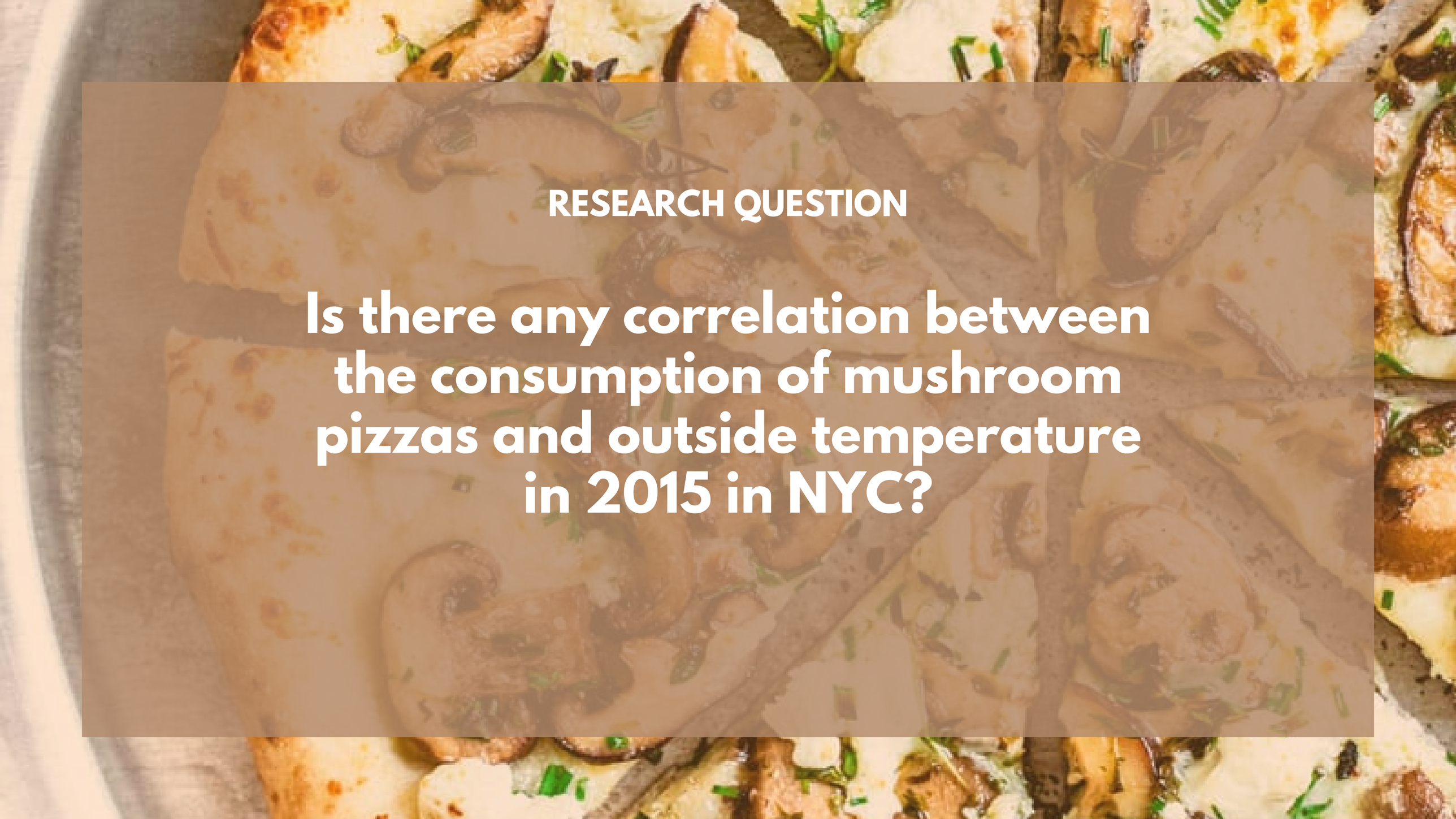
DEC 10TH 2022

TERM PROJECT 2

by Team 3
Caroline Hamberger, Chun-Hua Hsu, Hanna Asipovich



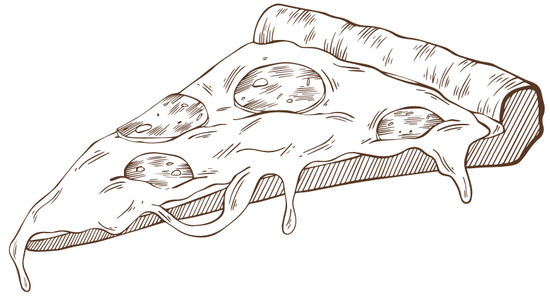
RESEARCH QUESTION

Is there any correlation between the consumption of mushroom pizzas and outside temperature in 2015 in NYC?

IMPORT PIZZA DATA

Using SQL to get an overview of one of our data sets from Term Project 1 - pizza sales data from NYC in 2015.

RESEARCH QUESTION & TECHNICAL APPROACH



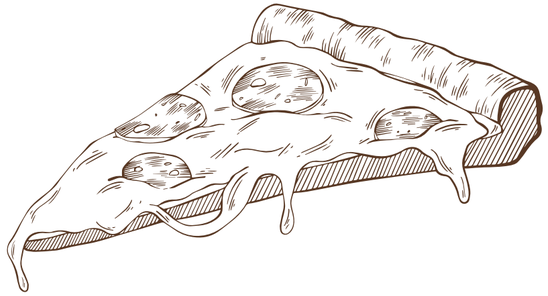
IMPORT PIZZA DATA

Using SQL to get an overview of one of our data sets from Term Project 1 - pizza sales data from NYC in 2015.

MATCH WEATHER DATA

Finding corresponding weather data (NYC in 2015, temperature in F°) online. Using postman to get it ready for KNIME.

RESEARCH QUESTION & TECHNICAL APPROACH



IMPORT PIZZA DATA

Using SQL to get an overview of one of our data sets from Term Project 1 - pizza sales data from NYC in 2015.

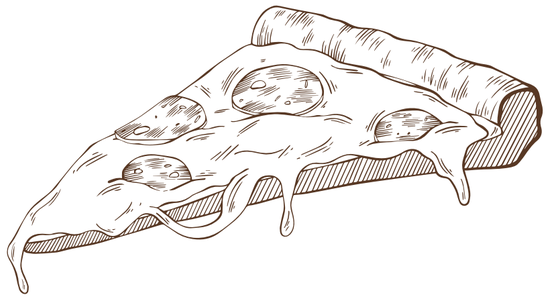
MATCH WEATHER DATA

Finding corresponding weather data (NYC in 2015, temperature in F°) online. Using postman to get it ready for KNIME.

IMPORT BOTH INTO KNIME

Creating two nodes: both a MySQL Connector/DB Query Reader, as well as String Manipulation/GET Request.

RESEARCH QUESTION & TECHNICAL APPROACH



IMPORT PIZZA DATA

Using SQL to get an overview of one of our data sets from Term Project 1 - pizza sales data from NYC in 2015.

MATCH WEATHER DATA

Finding corresponding weather data (NYC in 2015, temperature in F°) online. Using postman to get it ready for KNIME.

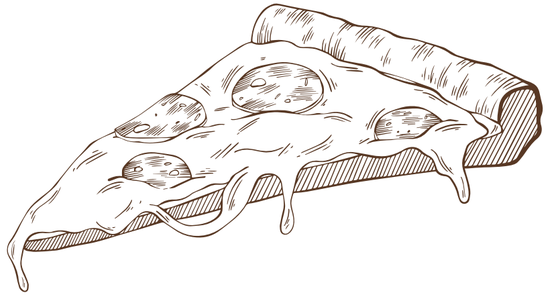
IMPORT BOTH INTO KNIME

Creating two nodes: both a MySQL Connector/DB Query Reader, as well as String Manipulation/GET Request.

CREATE A WORKLOW

Converting our weather data into a JSON format and producing aggregate table with SQL dataset, filtering relevant columns, and getting them ready for regression.

RESEARCH QUESTION & TECHNICAL APPROACH



IMPORT PIZZA DATA

Using SQL to get an overview of one of our data sets from Term Project 1 - pizza sales data from NYC in 2015.

MATCH WEATHER DATA

Finding corresponding weather data (NYC in 2015, temperature in F°) online. Using postman to get it ready for KNIME.

IMPORT BOTH INTO KNIME

Creating two nodes: both a MySQL Connector/DB Query Reader, as well as String Manipulation/GET Request.

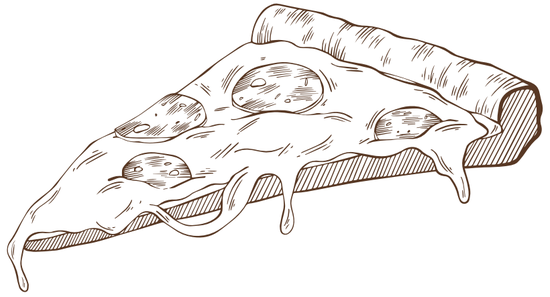
CREATE A WORKLOW

Converting our weather data into a JSON format and producing aggregate table with SQL dataset, filtering relevant columns, and getting them ready for regression.

REGRESSION OUTPUT

Using a regression learner and predicter to create a statistical table and line plot to analyze our data.

RESEARCH QUESTION & TECHNICAL APPROACH



DATA WAREHOUSE
FROM MYSQL

pizza

+

pizza_types

+

orders

+

order_details

ETL BY
KNIME

+

weather

SECOND DATASET
FROM API

DATA WAREHOUSE
FROM MYSQL

pizza

+

pizza_types

+

orders

+

order_details

+

weather

SECOND DATASET
FROM API

ETL BY
KNIME

ANALYTICAL LAYER

Sales volume of various pizzas on dates with max. temperature

DATA WAREHOUSE FROM MYSQL

pizza

+

pizza_types

+

orders

+

order_details

+

weather

SECOND DATASET FROM API

ETL BY KNIME

ANALYTICAL LAYER

Sales volume of various pizzas on dates with max. temperature

ETL BY KNIME

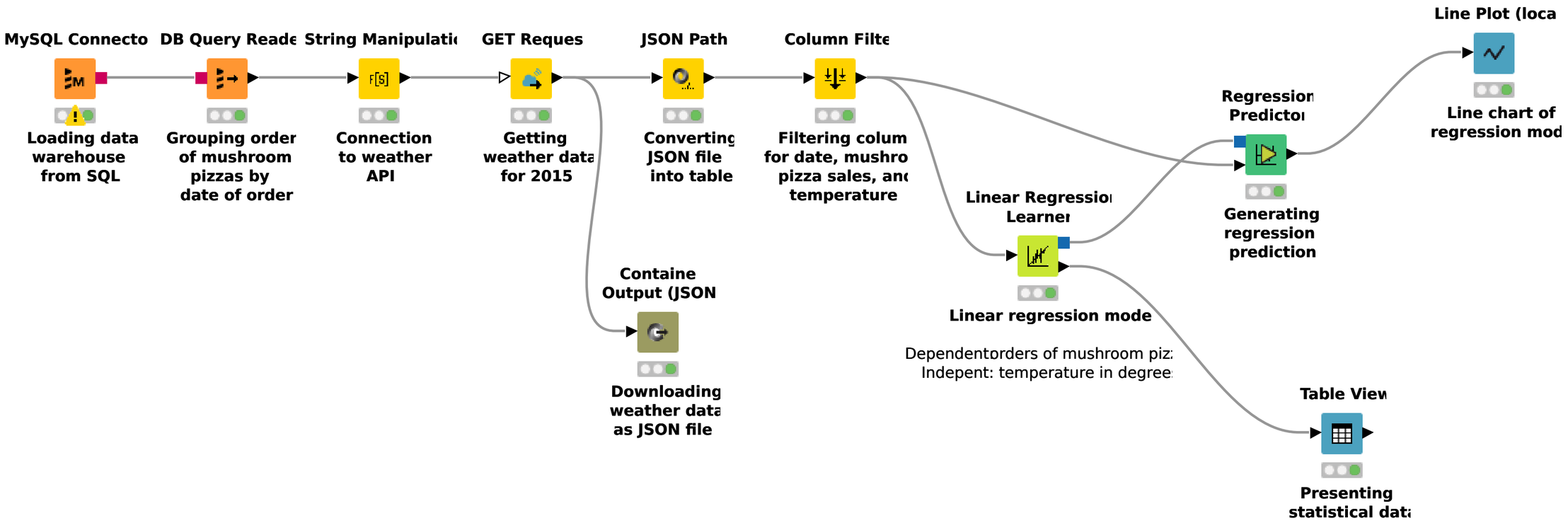
VISUALIZATION

Linear regression of temperature on mushroom pizza sales

Table view of linear regression

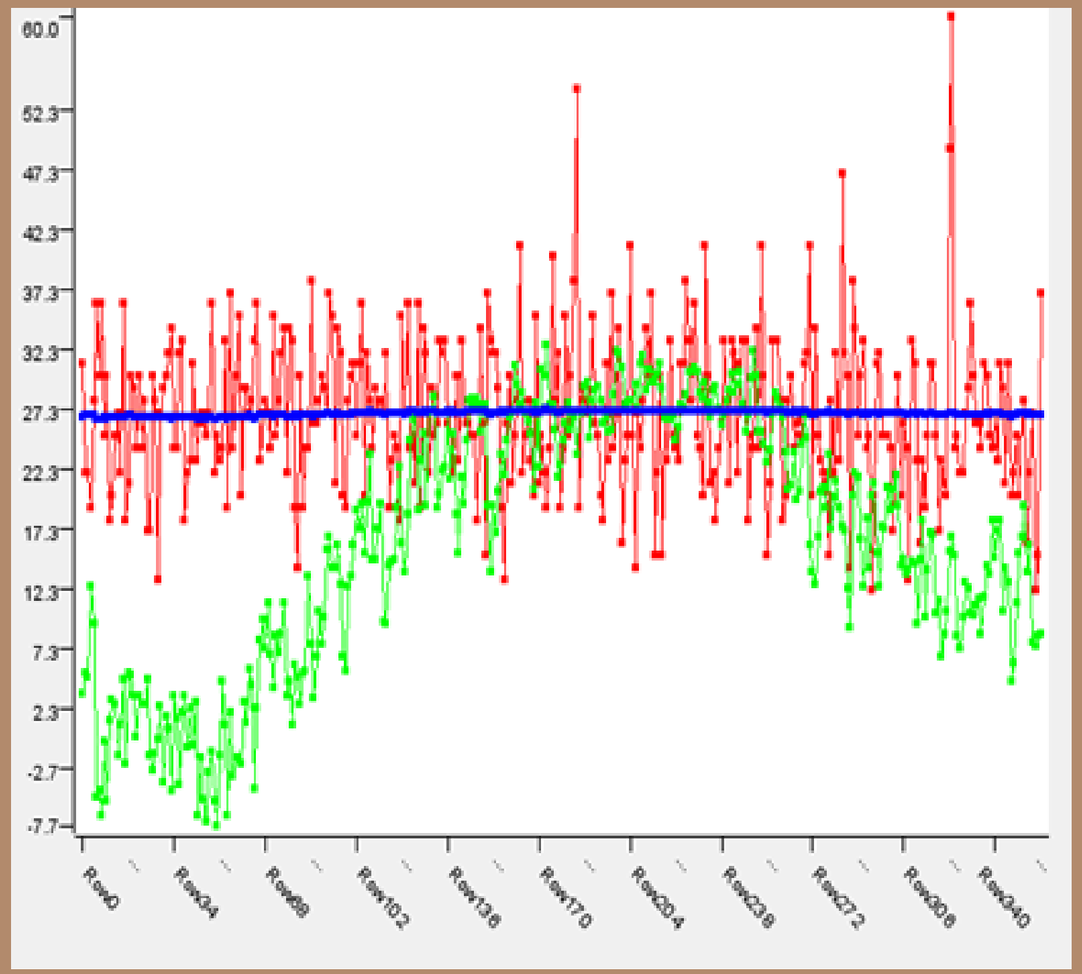
Regression prediction and line plot

KNIME WORKFLOW



REGRESSION OUTPUT

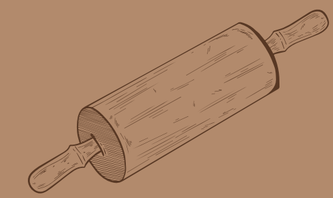
pizza orders = α + β * temperature + ε



Number of mushroom pizzas ordered

Temperature

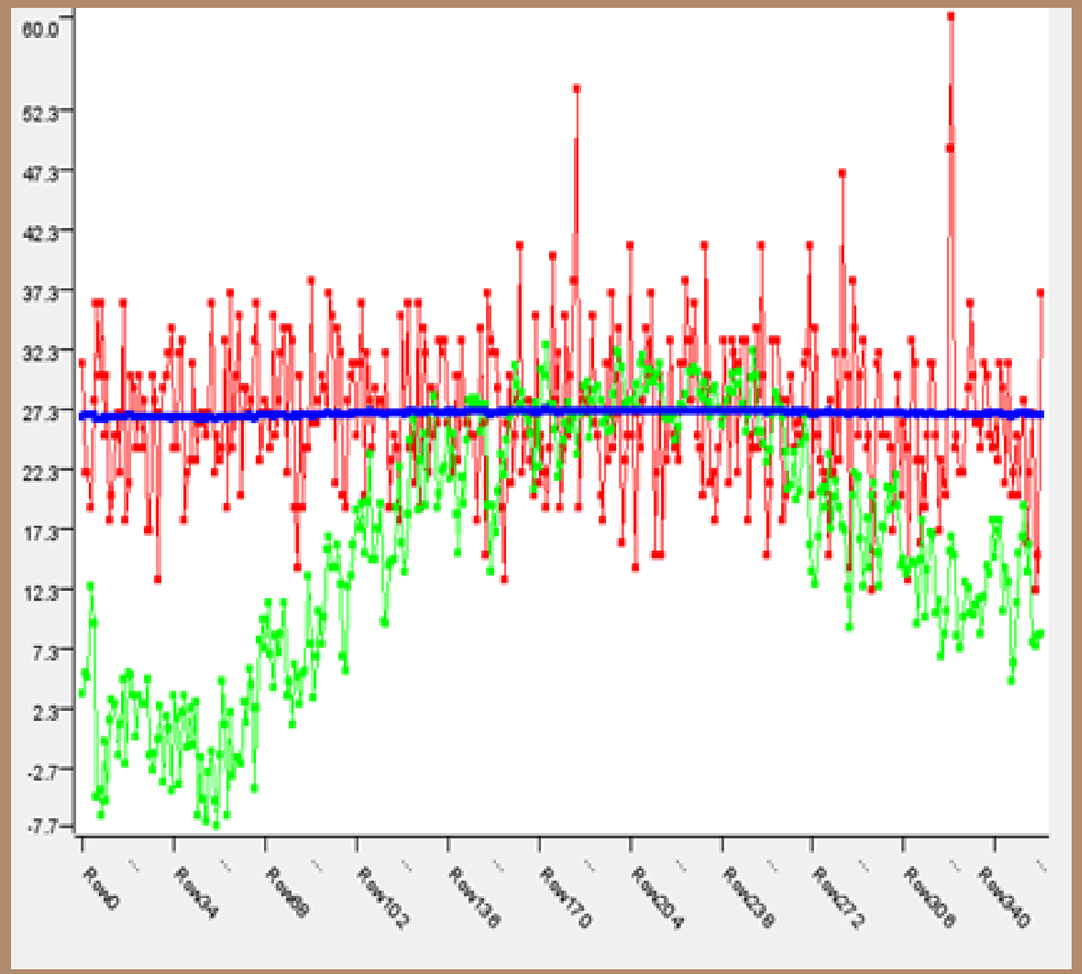
Regression prediciton



REGRESSION OUTPUT

pizza orders = α + β * temperature + ε

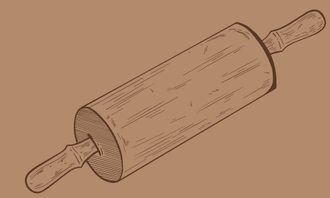
Visually, absolutely no correlation between the two lines. The blue prediction line indicates the same.



Number of mushroom pizzas ordered

Temperature

Regression prediciton

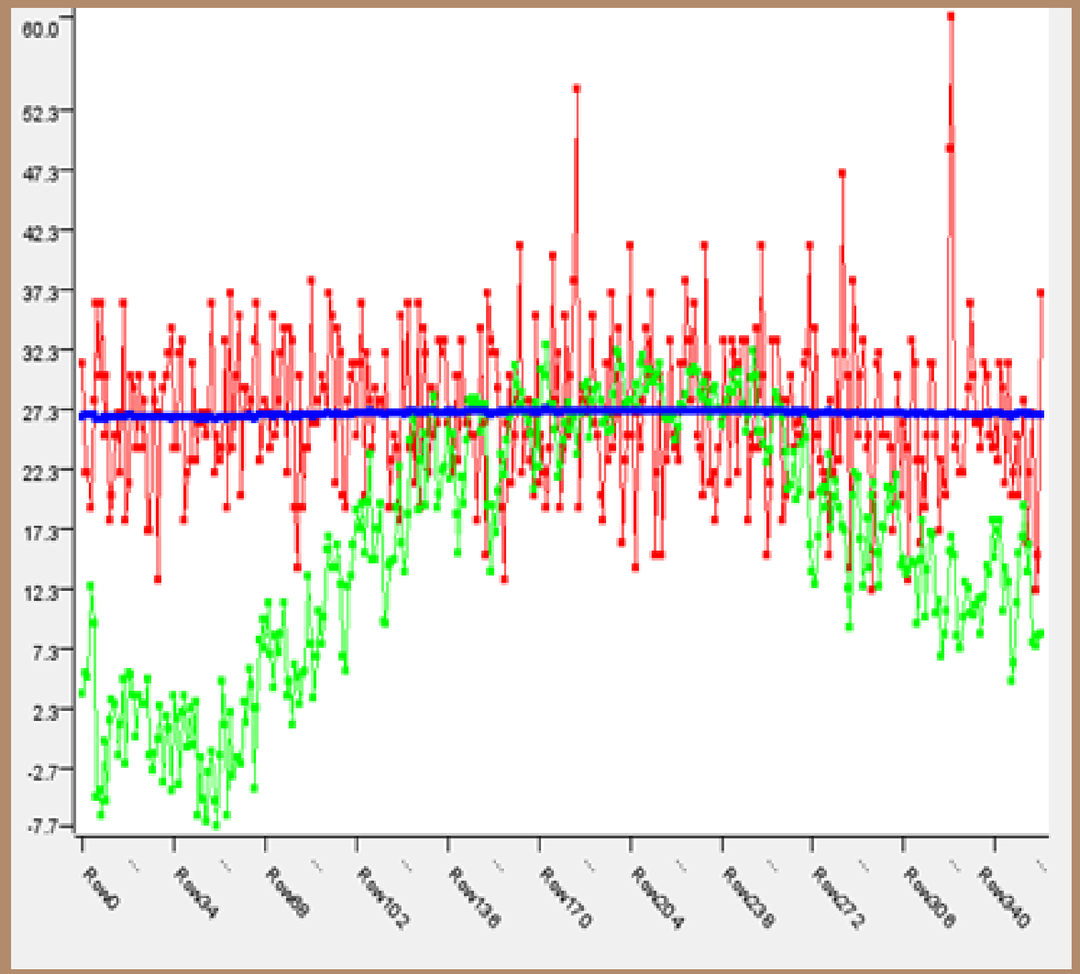


REGRESSION OUTPUT

pizza orders = α + β * temperature + ε

Visually, absolutely no correlation between the two lines. The blue prediction line indicates the same.

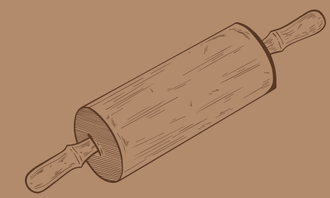
β coefficient: 0.019
Standard Error: 0.033



Number of mushroom pizzas ordered

Temperature

Regression prediciton



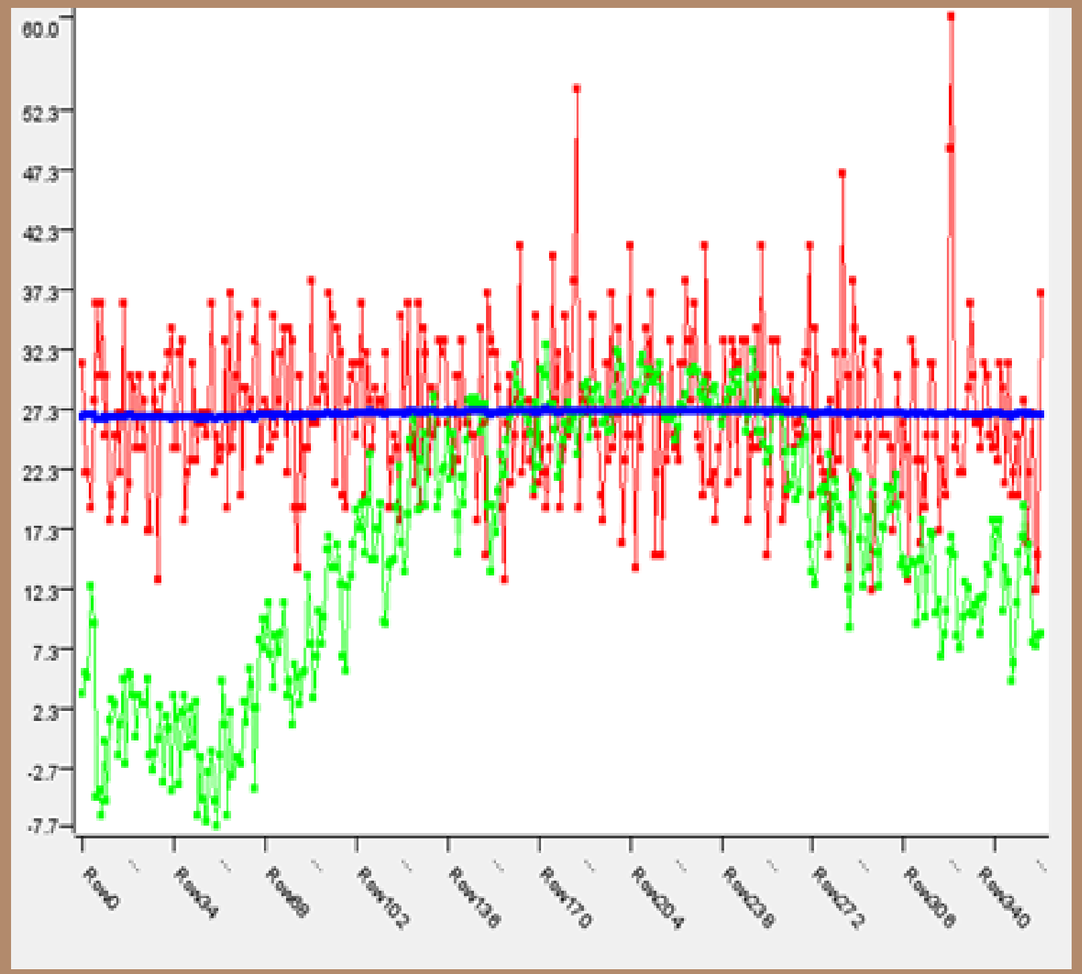
REGRESSION OUTPUT

pizza orders = α + β * temperature + ε

Visually, absolutely no correlation between the two lines. The blue prediction line indicates the same.

β coefficient: 0.019
Standard Error: 0.033

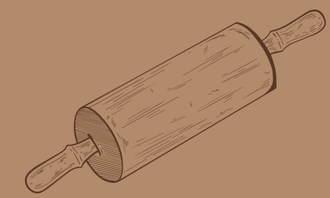
Indicates that there indeed is no statistically significant correlation.



Number of mushroom pizzas ordered

Temperature

Regression prediciton



DEC 10TH 2022

DON'T HESITATE TO ASK US ANY QUESTIONS!