Project Overview

This project is the first part of a two-part series. In the first part, you will blend and format data and deal with outliers.

For the second part, you will use your cleaned up dataset to create another linear regression model. The difference this time is that you will have to choose which variable(s) are the most important for the model using new techniques learned in the Selecting Predictor Variables section.

Scenario

Pawdacity is a leading pet store chain in Wyoming with 13 stores throughout the state. This year, Pawdacity would like to expand and open a 14th store. Your manager has asked you to perform an analysis to recommend the city for Pawdacity's newest store, based on predicted yearly sales.

How Do I Complete this Project?

This project uses skills learned throughout the "Data Preparation" lessons. To complete this project:

- Go through the course.
- Apply the skills learned in the course to solve the business problem given in the project details.
- Use our guidelines and rubric to help build your project.
- When you're ready, submit it to us for review using the submission template found in the supporting materials section.

Skills Required

In order to complete this project, you must be able to:

- **Understand different data types**
- **Deal with a variety of data issues**
- **Format data appropriately**
- **Blend data together using joins and unions**

NEXT