



Caterpillar Tube Pricing

Model quoted prices for industrial tube assemblies

Data Description

The dataset is comprised of a large number of relational tables that describe the physical properties of tube assemblies. You are challenged to combine the characteristics of each tube assembly with supplier pricing dynamics in order to forecast a quote price for each tube. The quote price is labeled as *cost* in the data.

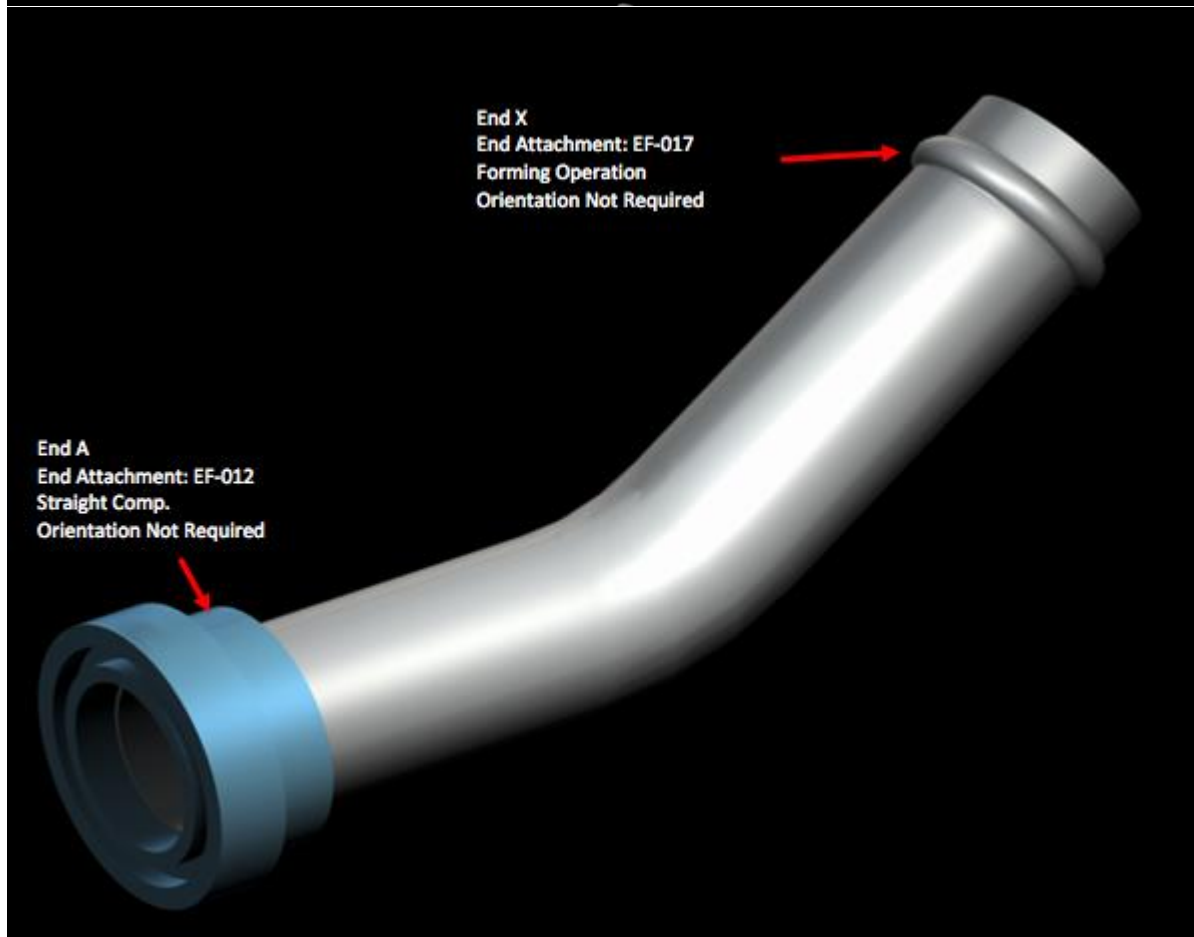
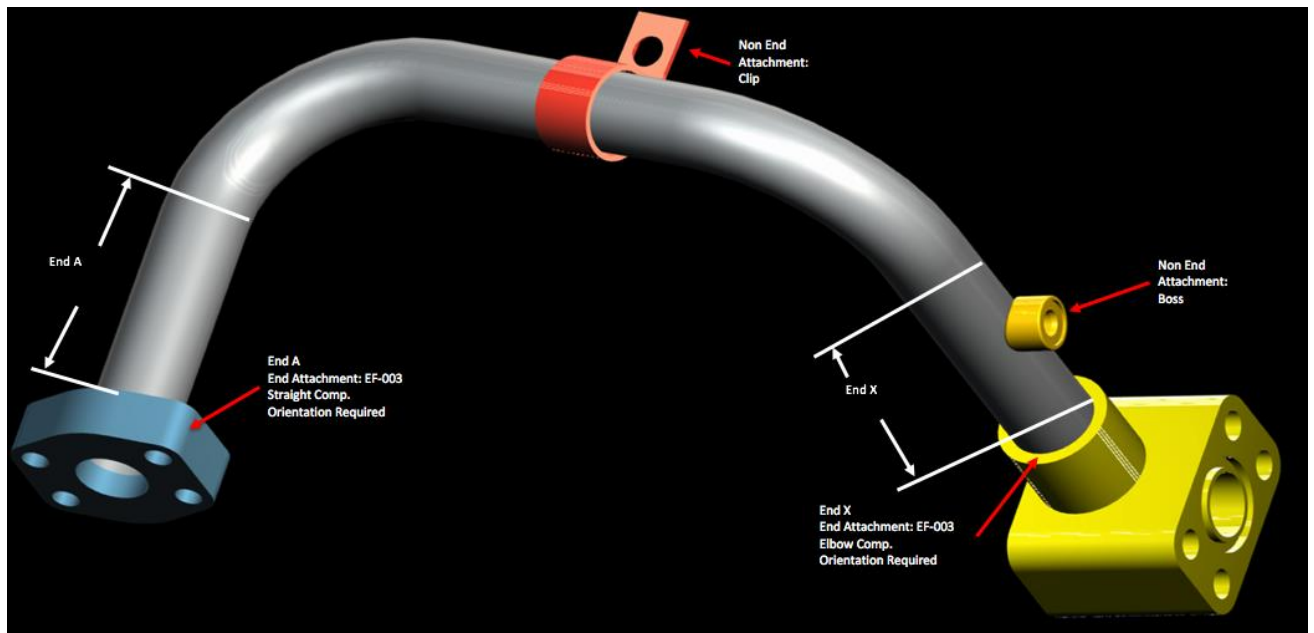
File descriptions

train_set.csv and test_set.csv

This file contains information on price quotes from our suppliers. Prices can be quoted in 2 ways: bracket and non-bracket pricing. Bracket pricing has multiple levels of purchase based on *quantity* (in other words, the cost is given assuming a purchase of *quantity* tubes). Non-bracket pricing has a minimum order amount (*min_order*) for which the price would apply. Each quote is issued with an *annual_usage*, an estimate of how many tube assemblies will be purchased in a given year.

tube.csv

This file contains information on tube assemblies, which are the primary focus of the competition. Tube Assemblies are made of multiple parts. The main piece is the tube which has a specific diameter, wall thickness, length, number of bends and bend radius. Either end of the tube (End A or End X) typically has some form of end connection allowing the tube assembly to attach to other features. Special tooling is typically required for short end straight lengths (*end_a_1x*, *end_a_2x* refer to if the end length is less than 1 times or 2 times the tube diameter, respectively). Other components can be permanently attached to a tube such as bosses, brackets or other custom features.



bill_of_materials.csv

This file contains the list of components, and their quantities, used on each tube assembly.

specs.csv

This file contains the list of unique specifications for the tube assembly. These can refer to materials, processes, rust protection, etc.

tube_end_form.csv

Some end types are physically formed utilizing only the wall of the tube. These are listed here.

components.csv

This file contains the list of all of the components used. Component_type_id refers to the category that each component falls under.

comp_[type].csv

These files contain the information for each component.

type_[type].csv

These files contain the names for each feature.