Intelligent Data Analytics

Charles Nicholson, PhD

School of Industrial and Systems Engineering

University of Oklahoma

OKC Innovates Symposium: Big Data
October 18, 2017 – Oklahoma City, OK

Who was Tycho Brahe?

(1546 – 1601)

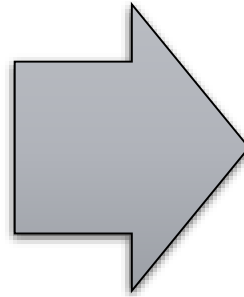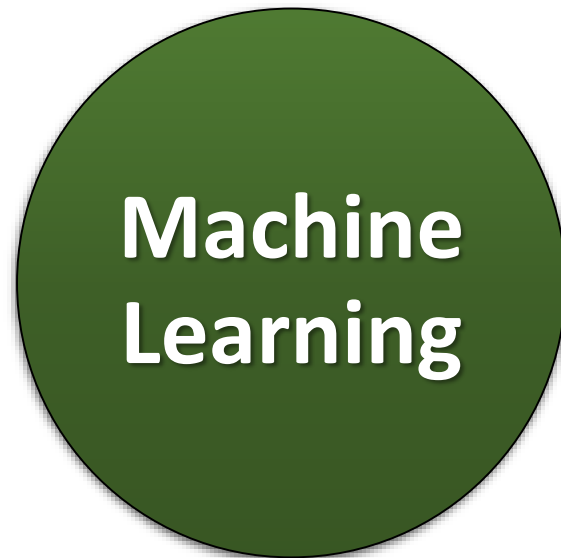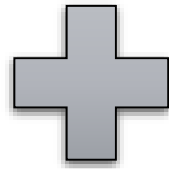# Who was Johannes Kepler?

(1571 – 1630)

- Brahe: 16th century astronomer who collected and cataloged a vast amount of measurements of the positions of the sun, moon, and planets, i.e., the "big data" of his day.
- However, he was unsuccessful in finding a consistent scheme to tie all of it together.
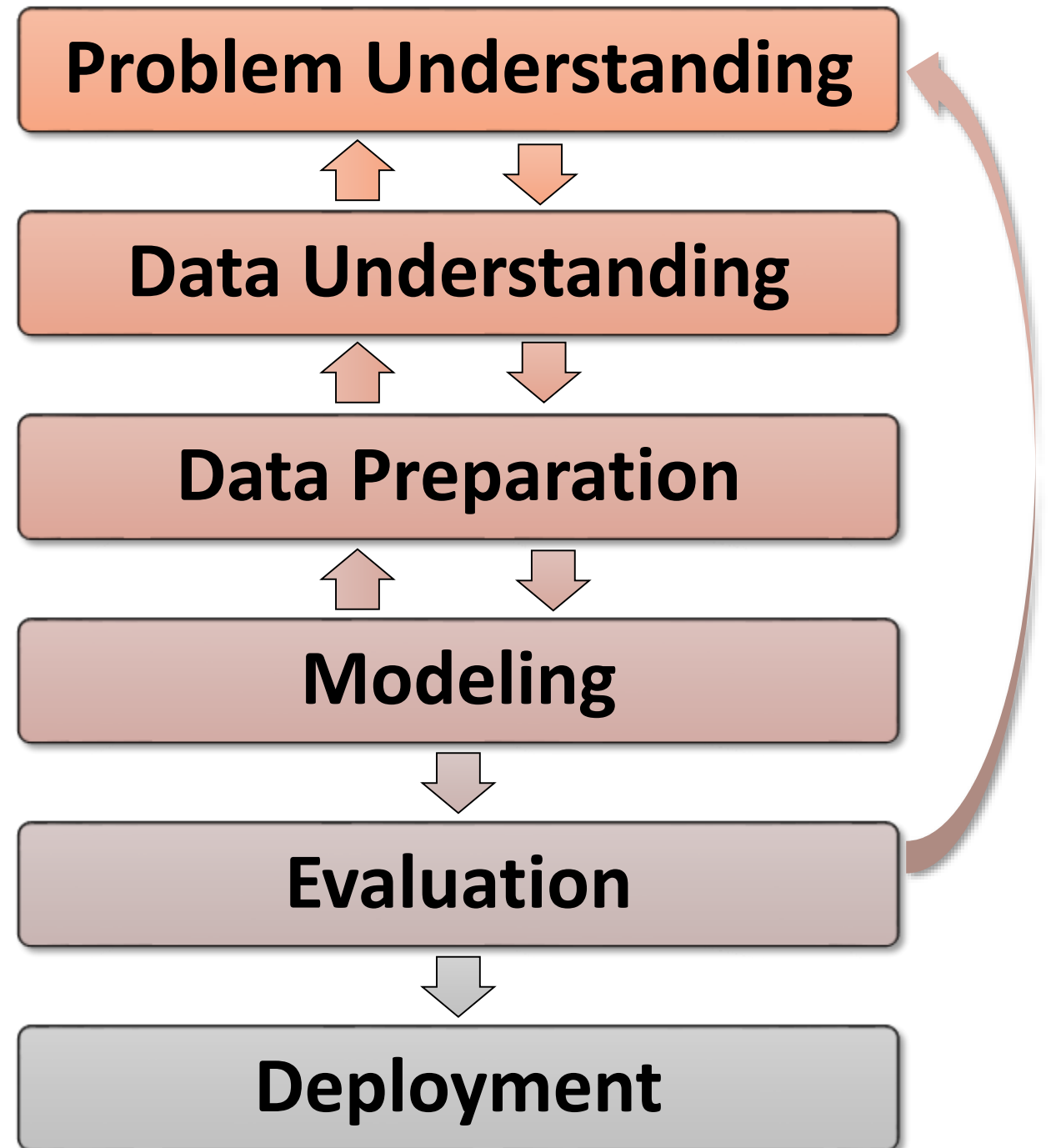- Today his work is almost forgotten.

- Kepler was Tycho's assistant.
- He used the available "big data" and, together with his genius, developed something revolutionary.
- His laws of planetary motion are still taught in Physics classes 400 years later!

Data alone are not enough.

Human Problem Solving + Machine Learning → Intelligent Data Analytics

CRISP-DM

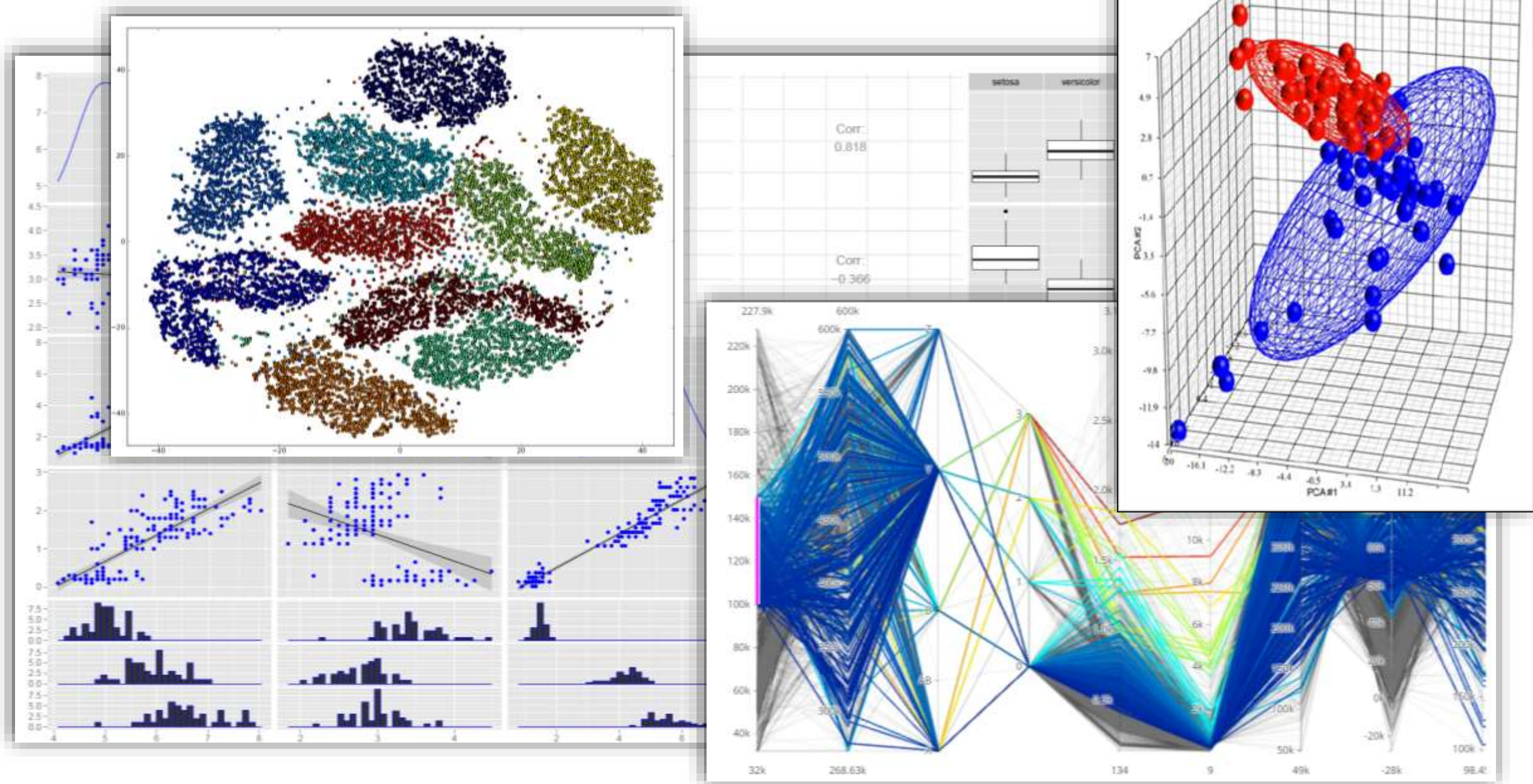Cross Industry Standard Process for Data Mining
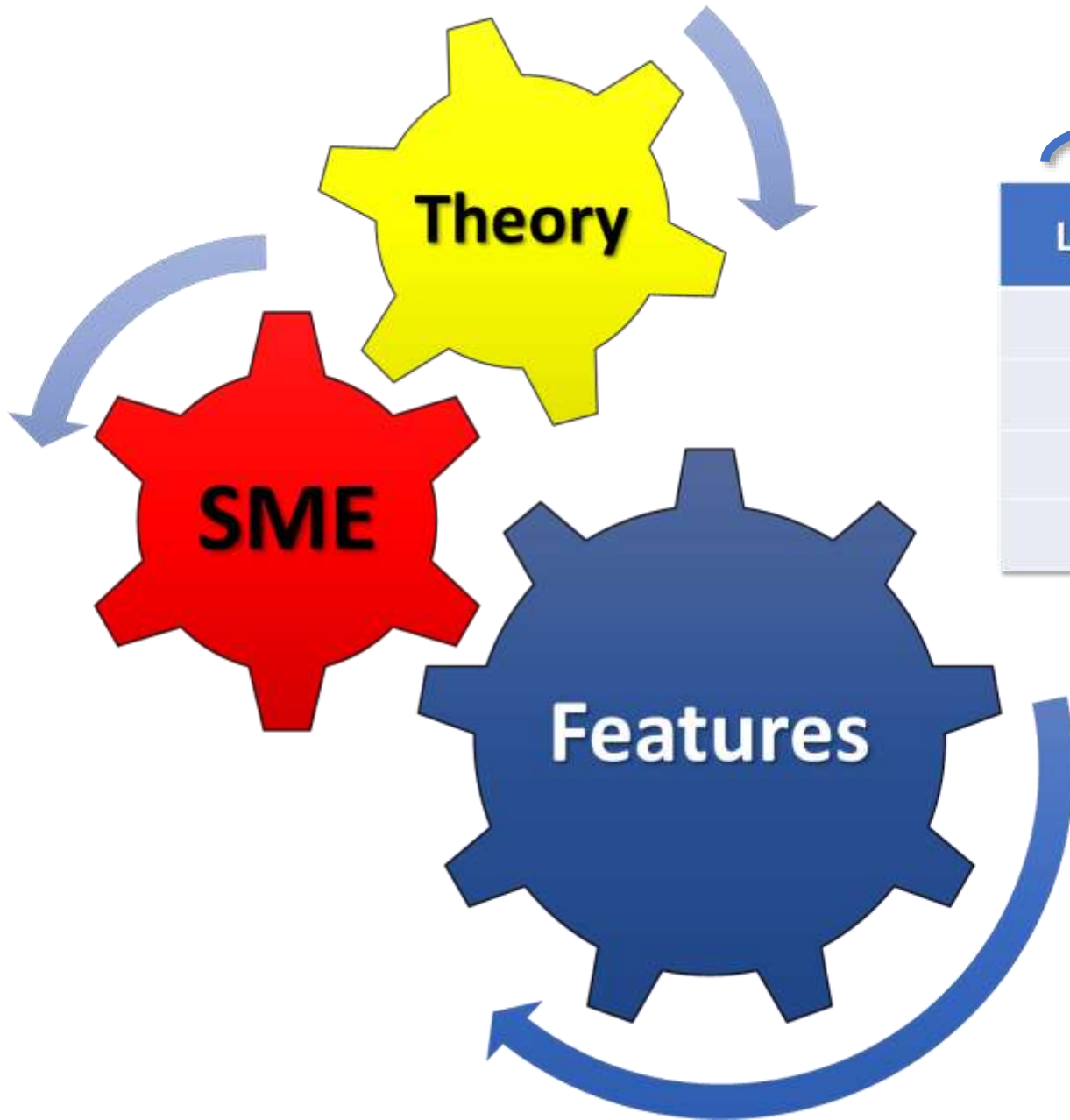
I: Understanding the Problem

A problem well stated is a problem half solved.

-Charles Kettering

II: Exploratory Data Analysis (EDA)
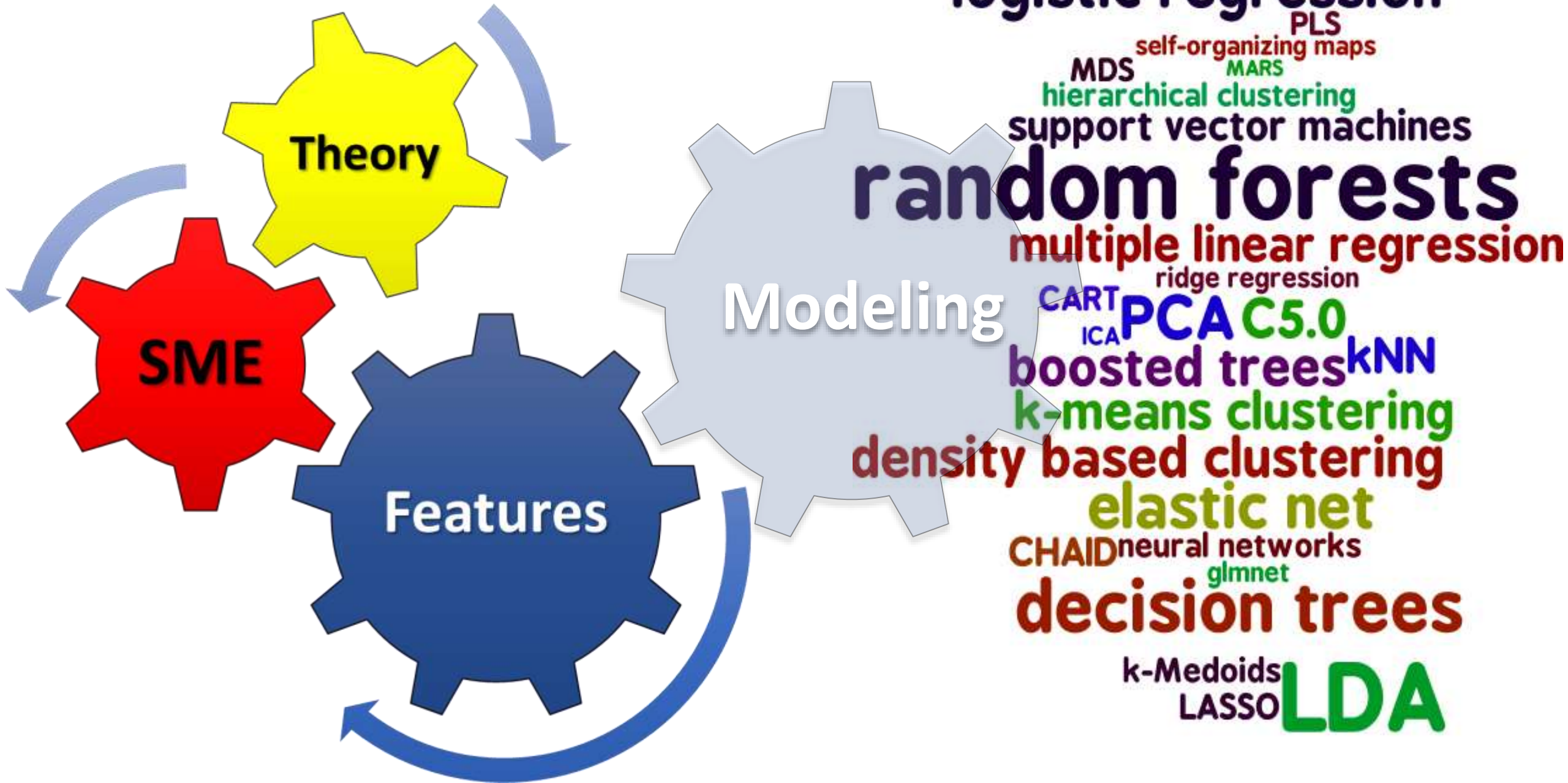
III: Feature Engineering

input features

Latitude	Longitude	Latitude	Longitude	Distance
San Juan, PR		New York, NY		1609
Oklahoma City, OK		Beijing		6807
Moscow		Washington D.C.		4865
Vancouver, BC		Tulsa, OK		1637

$$\Delta\sigma = \arctan \frac{\sqrt{(\cos\phi_2 \cdot \sin(\Delta\lambda))^2 + (\cos\phi_1 \cdot \sin\phi_2 - \sin\phi_1 \cdot \cos\phi_2 \cdot \cos(\Delta\lambda))^2}}{\sin\phi_1 \cdot \sin\phi_2 + \cos\phi_1 \cdot \cos\phi_2 \cdot \cos(\Delta\lambda)}$$

Theory

SME

Features

III: Feature Engineering

IV: Delivering Insights

"Make sure a candidate can find a story in a data set and provide a coherent narrative about a key data insight... "

—— Harvard Business Review

Fi you can raed this, yuo are pbraloby hmuan. Yuor mnid deos not raed ervey lteter by istlef, but rcgoeznies the wrod as a wlhoe. Nveer froget the pweor and roseurcfelunses of the hmaun mnid to rcegnozie ptatrens, slove pbrolmes, and to ionvnate!