Introduction to Data Science

Last Update: 2 March 2020

Chapter 5
Prescriptive Analysis

Papangkorn Inkeaw, PhD

Department of Computer Science, Faculty of Science

Chiang Mai University

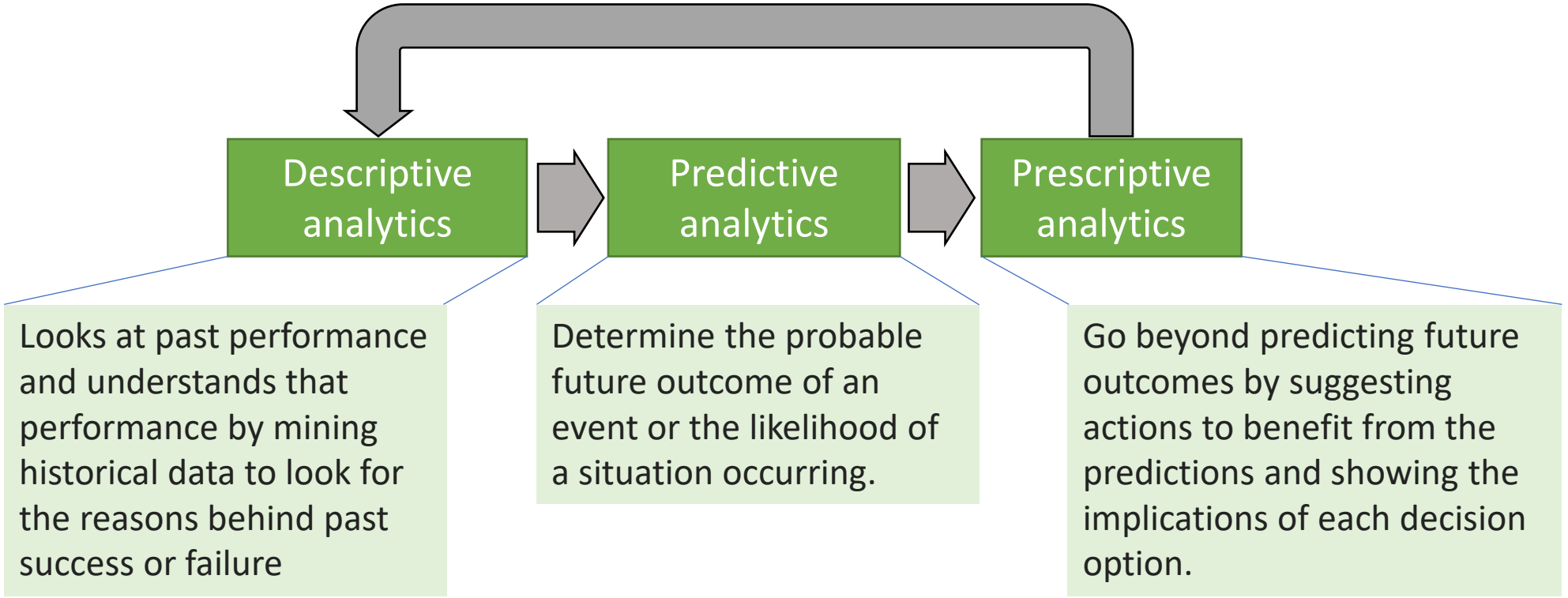
Outline

Prescriptive Analysis

1. **Prescriptive Analysis**
2. **Optimization**
    - **Optimization Problem**
    - **Defining an optimization problem**
    - **Applying predictive analytics to solve optimization problem**
3. **Simulation**
    - **What is simulation?**
    - **Applying Simulation to Prescriptive Analysis**
    - **Monte Carlo simulation**

Prescriptive Analysis

Process of Business Analytics



Prescriptive Analysis

- **Prescriptive analytics** focuses on finding the best course of action in a scenario given the available data.
- It use **optimization** and **simulation** to advise on possible outcomes.
- It leverages predictive analytics and descriptive analytics to derive ideal outcomes or solutions.

Optimization

Prescriptive Analysis

Optimization Problem

Find the best solution from all feasible solutions

The solution that achieve an objective

Minimize or maximize an objective function

All possible solution under constrains

Optimization

Prescriptive Analysis

Optimization

How to define an optimization problem.

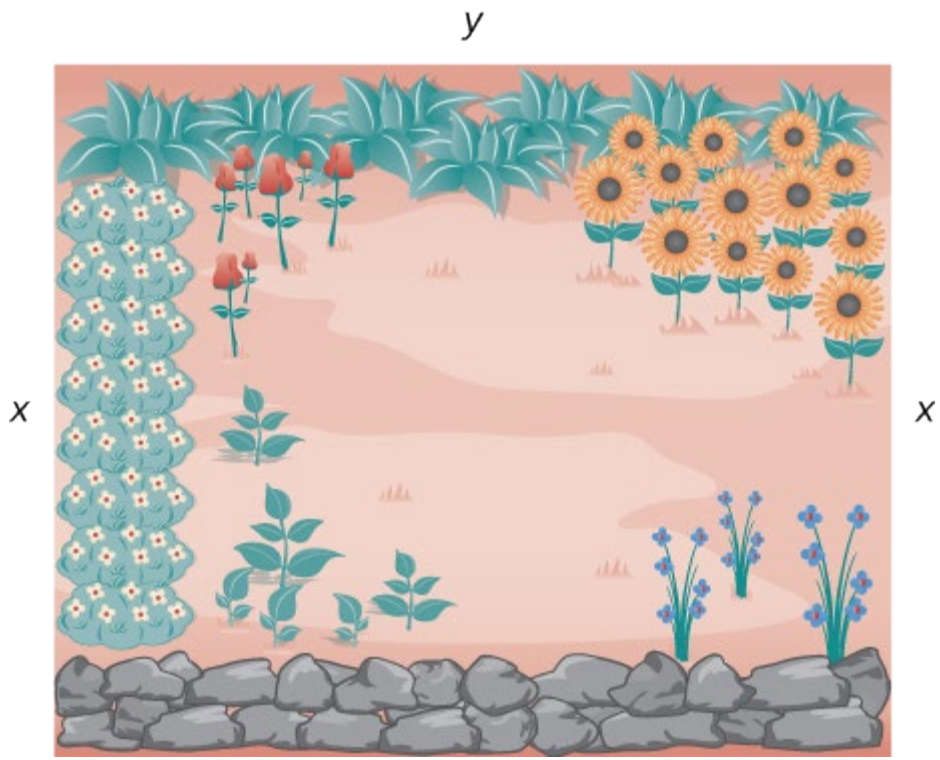
- **What is the objective/goal?**
- **What are constrains (on possible solution)?**

Optimization

Prescriptive Analysis

Simple Example

A rectangular garden is to be constructed using a rock wall as one side of the garden and wire fencing for the other three sides Given 100 ft of wire fencing, determine the dimensions that would create a garden of maximum area. What is the maximum area?



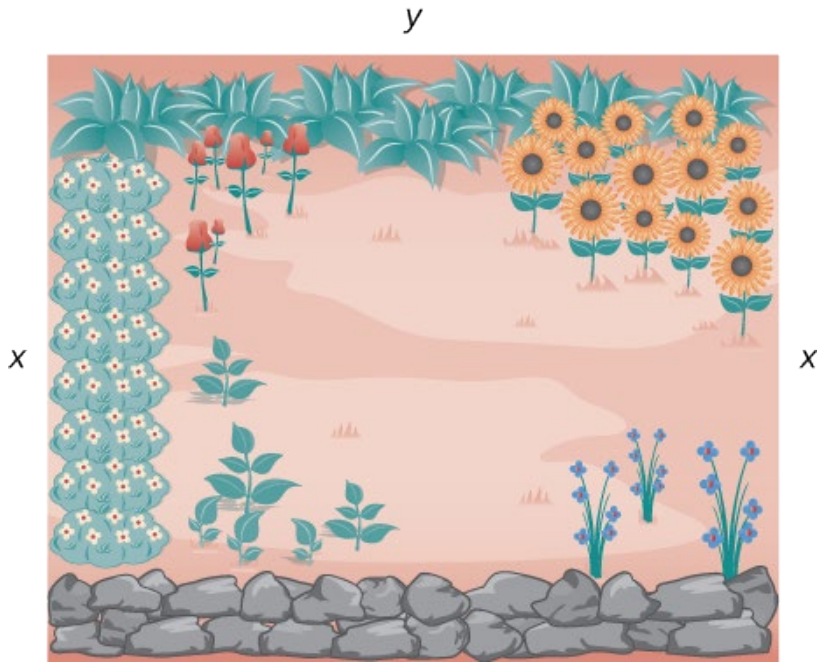
Source:
https://math.libretexts.org/Bookshelves/Calculus/Map%3A_Calculus_-_Early_Transcendentals_(Stewart)/04%3A_Applications_of_Differentiation/4.07%3A_Optimization_Problems

Optimization

Prescriptive Analysis

Simple Example

A rectangular garden is to be constructed using a rock wall as one side of the garden and wire fencing for the other three sides Given 100 ft of wire fencing, determine the dimensions that would create a garden of maximum area. What is the maximum area?



- **What is the objective/goal?**

  Find the dimensions that maximize the area.

  $area = xy$

- **What are constrains (on possible solution)?**

  $2x + y \leq 100$

Optimization

Prescriptive Analysis

Quiz

A plant produces and sells semiconductor devices. The cost per one unit (also known as the unit cost) depends on the volume of production and includes a fixed part 1000 ($/*device*) and a variable part 2n ($/*device*), where n is the number of units produced per month. The price of the device, in turn, depends on the volume of production according to the law p(n)=10000−n ($/*device*). Determine at what volume of production the profit will be highest?

What is the objective/goal?

Source: https://www.math24.net/optimization-problems-economics/#example4

Optimization

Prescriptive Analysis

How to solve an optimization problem.

The objectives can be defined as math formulas

Apply mathematical optimization techniques such as:

- Mathematical programming
- Evolutionary computation

The objectives cannot be defined as math formulas, but you have historical data

Apply data-driven techniques such as:

- classification
- regression

Note:

- **For complicate problem, some cases, the objective cannot be defined as math formulas.**
- **Historical/transection data can be used to overcome the problem.**

Optimization

Prescriptive Analysis

Applying predictive analytics to prescriptive analytics

Example: Broken pipes and prescriptive analytics

Each sensor sent data to the operational database at all times

Operational Database

Company will be able to see the status of its network in real time.

Optimization

Prescriptive Analysis

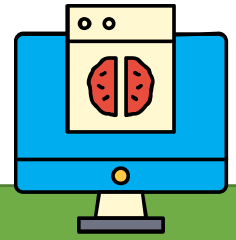
Applying predictive analytics to prescriptive analytics

Example : Broken pipes and prescriptive analytics

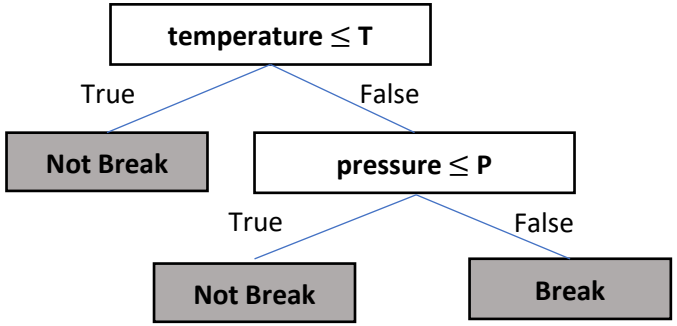
Database

D	X_1	X_2	Y
$\mathbf{x}_1$			
...			
$\mathbf{x}_n$			

pressure
temperature
broken pipe



Train a predictive model

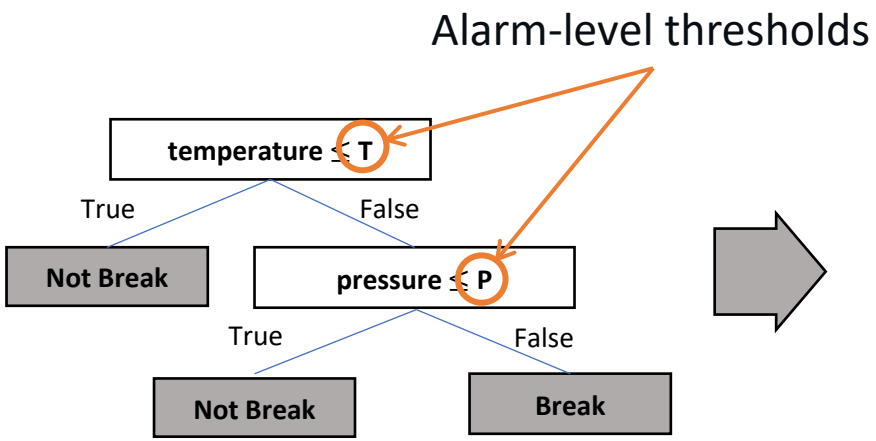


Predictive Analysis

Optimization
Prescriptive Analysis

Applying predictive analytics to prescriptive analytics

Example : Broken pipes and prescriptive analytics



Rules for Decision-making:

If atmospheric pressure and temperature are above alarm-level thresholds

Then

- lowering the water flow
- using alternative routes

Prescriptive Analysis

Source: https://www.zdnet.com/article/a-guide-for-prescriptive-analytics-the-art-and-science-of-choosing-and-applying-the-right-techniques/

Optimization

Prescriptive Analysis

Applying predictive analytics to prescriptive analytics

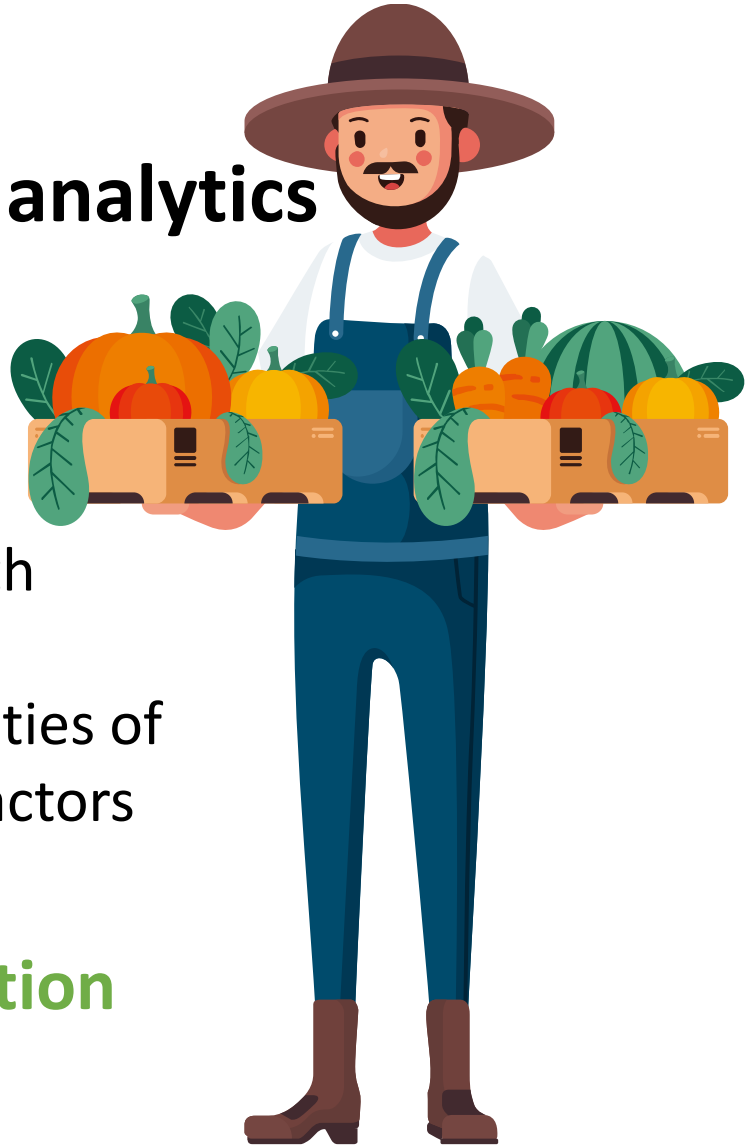
Thank about this simple problem

A farmer want to improve the production of avocados.
He has a historical dataset that contains

- the quantity of fertilizer providing to a tree for each month
- the quantity of the avocados produced in one season

In practice, we don’t know exact correlation between quantities of fertilizer and avocados. Moreover, there are unobservable factors of the production of avocados.

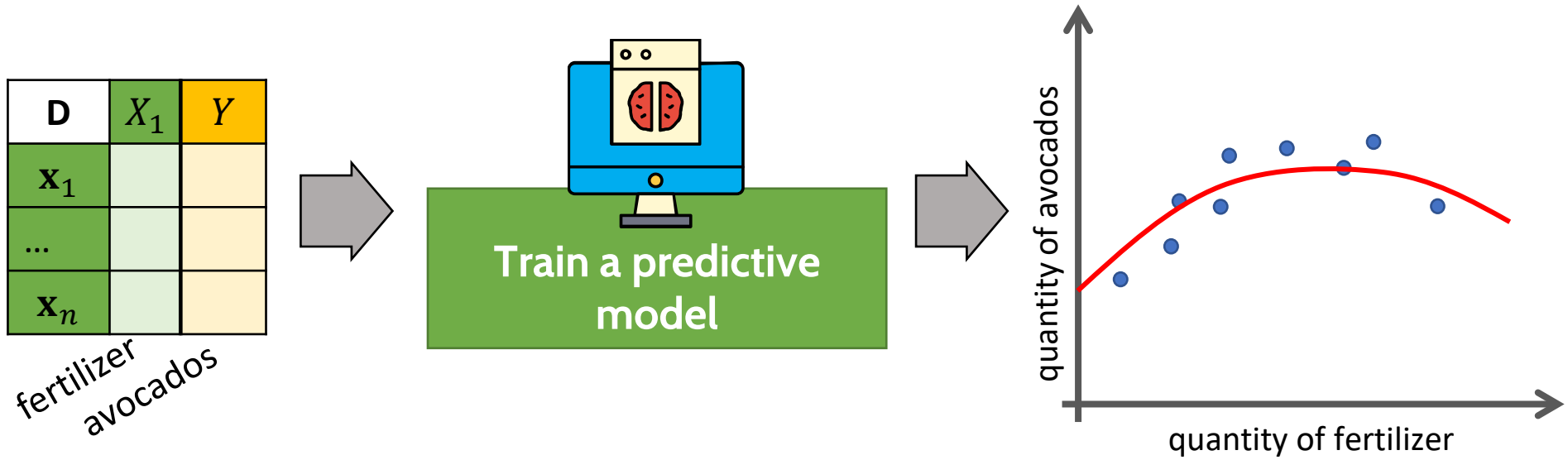
What should the farmer do to improve the production of avocados and minimize the cost of fertilizer?



Optimization

Prescriptive Analysis

Applying predictive analytics to prescriptive analytics

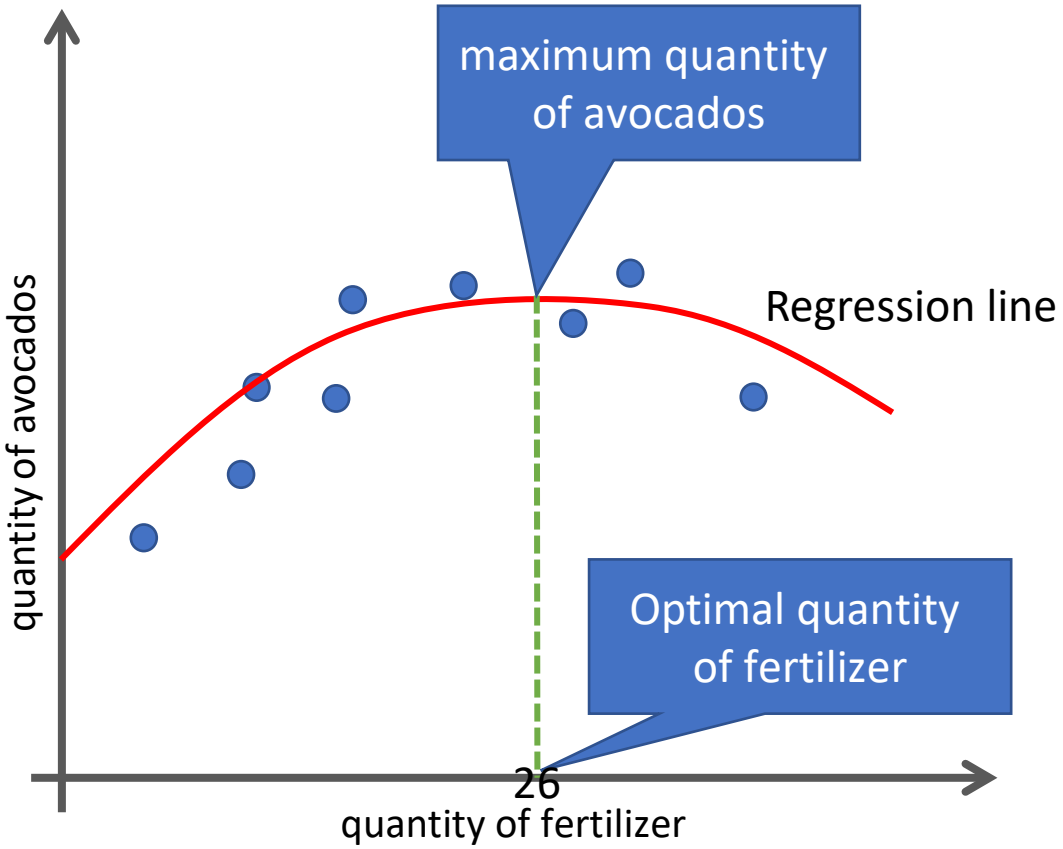


Predictive Analysis

Optimization

Prescriptive Analysis

Applying predictive analytics to prescriptive analytics



Prescriptive Analysis

Further Study

Papers:

- Katerina Lepenioti, Alexandros Bousdekis, Dimitris Apostolou, Gregoris Mentzas, Prescriptive analytics: Literature review and research challenges, International Journal of Information Management, Volume 50, 2020, Pages 57-70, https://doi.org/10.1016/j.ijinfomgt.2019.04.003.
- Lepenioti K., Bousdekis A., Apostolou D., Mentzas G. (2019) Prescriptive Analytics: A Survey of Approaches and Methods. In: Abramowicz W., Paschke A. (eds) Business Information Systems Workshops. BIS 2018. Lecture Notes in Business Information Processing, vol 339. Springer, Cham

Website

- https://www.zdnet.com/article/a-guide-for-prescriptive-analytics-the-art-and-science-of-choosing-and-applying-the-right-techniques/

Simulation

Prescriptive Analysis

The process of modelling a real-life or hypothetical situation on a computer.

Simulation

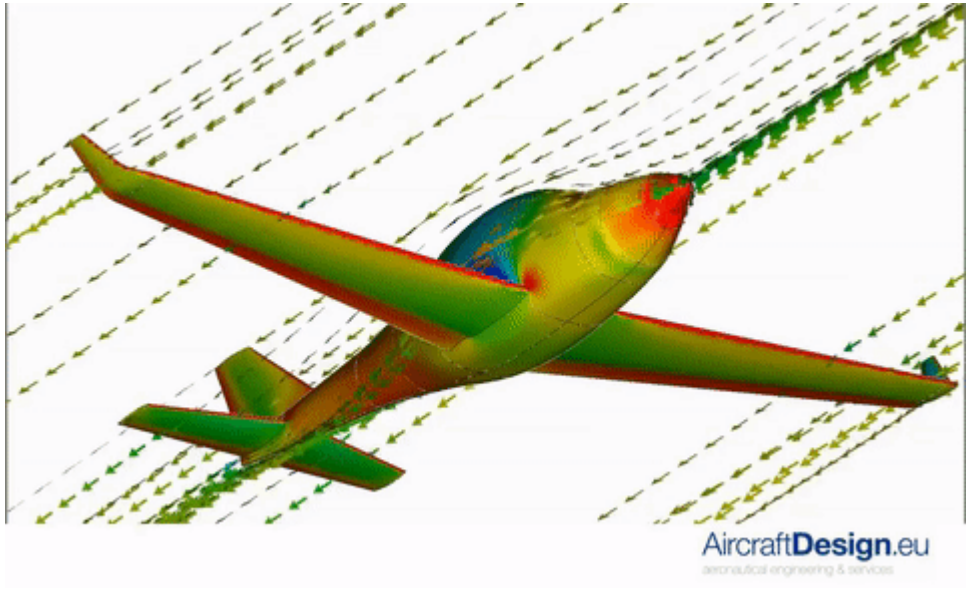
Prescriptive Analysis

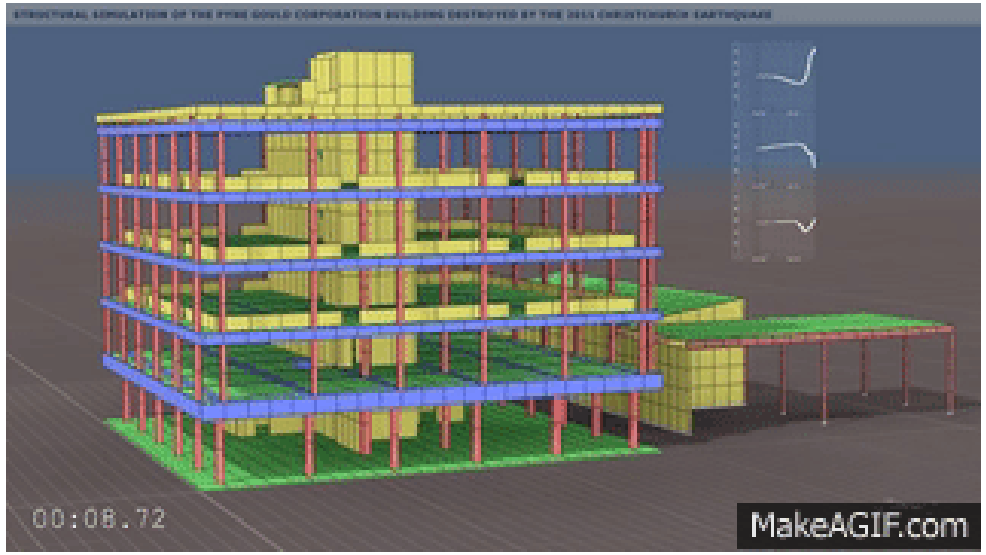
Simulation is useful because it can be used to test new ideas about business decisions and actions to mitigate risk pertaining to a process or system, or how modifications will affect an existing process or system.

Simulation

Prescriptive Analysis

Simulation is often used in applications related to safety of infrastructure as well as safety, quality, and design of products.





Simulation

Prescriptive Analysis

Applying Simulation to Prescriptive Analysis

Example: Optimum number of open Check-out counters

The aim is to experiment with different queuing systems to understand their customers better.

- For instance, there are 20 counters in total.
- Each open counter costs $20 to operate.
- Customers who wait more than 10 minutes are ‘annoyed’ and there’s a penalty of $1 per customer.

What is the optimal number of counters?

Simulation

Prescriptive Analysis

Applying Simulation to Prescriptive Analysis

Example: Optimum number of open Check-out counters

What is the optimal number of counters?

Option 1: Perform experiment of different numbers of counters on real situation.

Option 2: Simulate experiment of different numbers of counters on computer.

Simulation

Prescriptive Analysis

Applying Simulation to Prescriptive Analysis

Example: Optimum number of open Check-out counters

Assumption:

Each counter has its queue and arriving customer joins one queue randomly.

Random under a certain distribution

Gamma distributions are assumed for inter-arrival time and service time for customers.

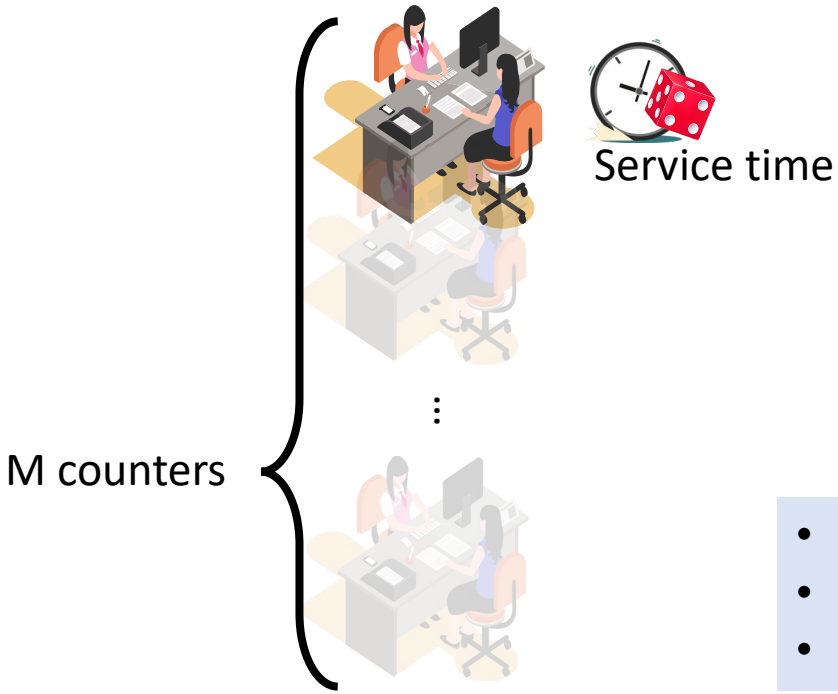
Simulation

Prescriptive Analysis

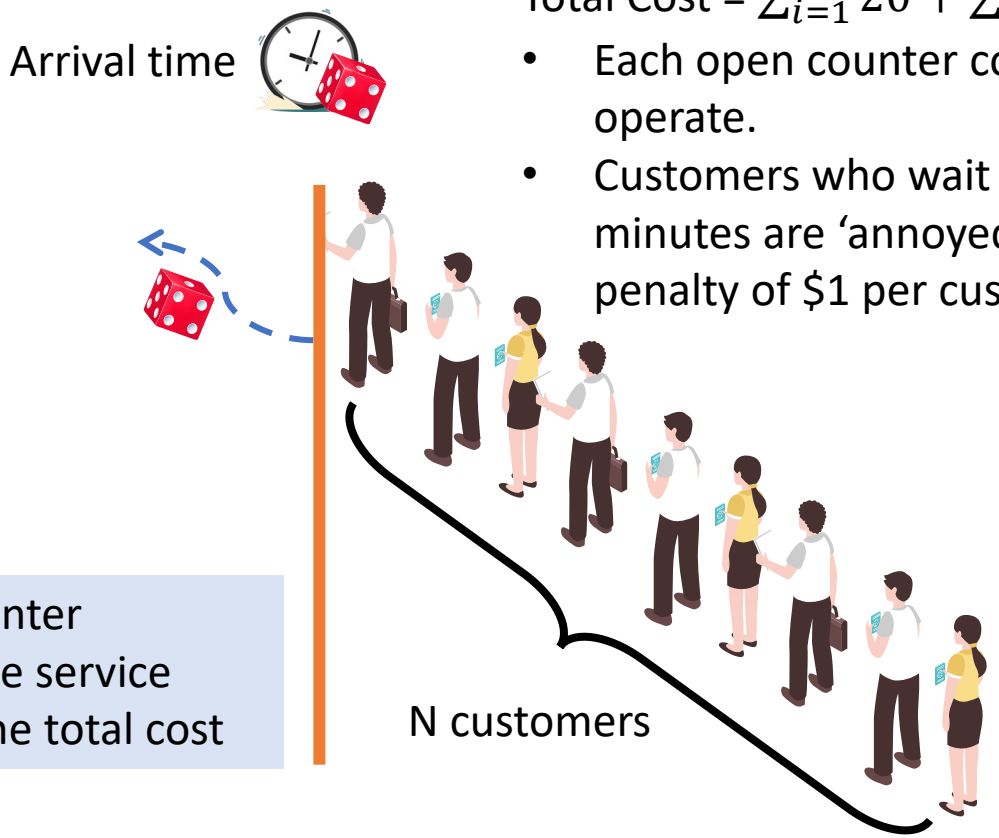
Applying Simulation to Prescriptive Analysis

Example: Optimum number of open Check-out counters

Queue Simulation



- Open 1 counter
- Simulate the service
- Calculate the total cost



Total Cost = $\sum_{i=1}^{M} 20 + \sum_{j=1}^{N} (t_j^{wait} > 10)$

- Each open counter costs $20 to operate.
- Customers who wait more than 10 minutes are ‘annoyed’ and there’s a penalty of $1 per customer.

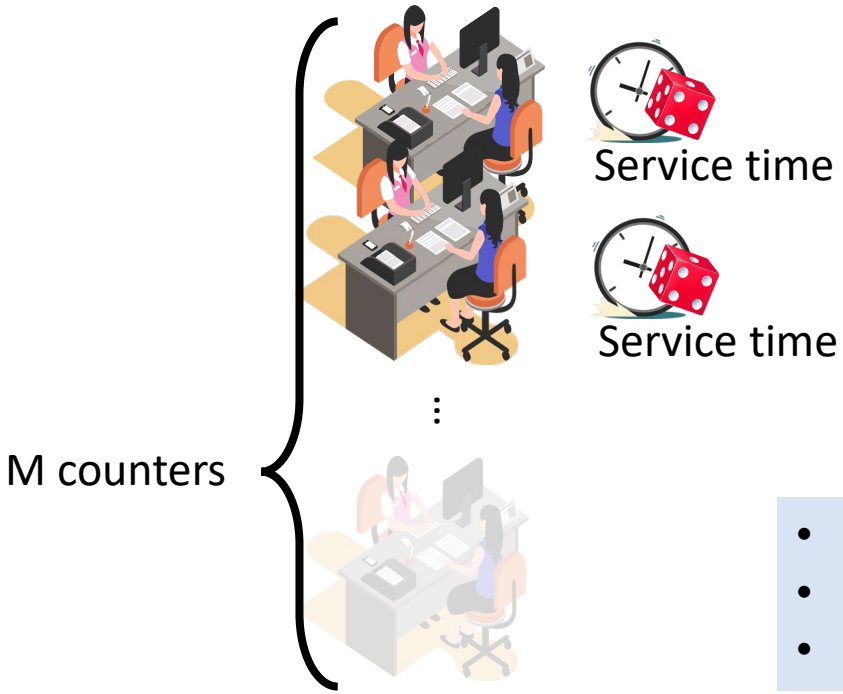
Simulation

Prescriptive Analysis

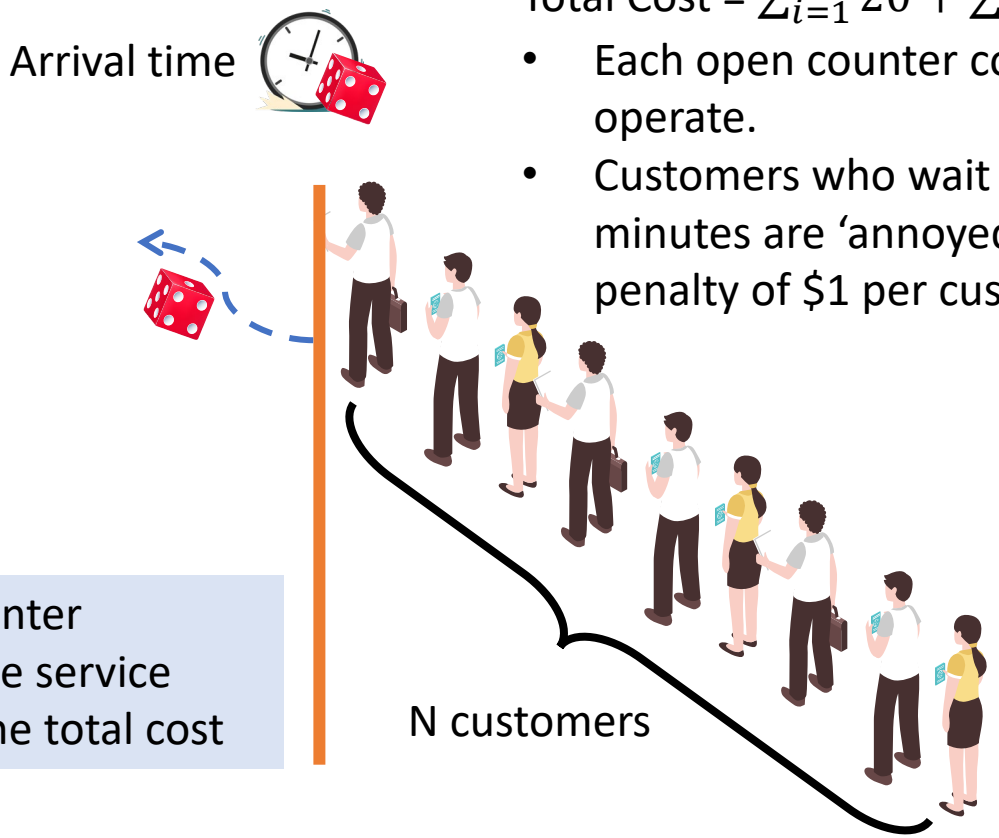
Applying Simulation to Prescriptive Analysis

Example: Optimum number of open Check-out counters

Queue Simulation



- Open 2 counter
- Simulate the service
- Calculate the total cost



Total Cost = $\sum_{i=1}^{M} 20 + \sum_{j=1}^{N} (t_j^{wait} > 10)$

- Each open counter costs $20 to operate.
- Customers who wait more than 10 minutes are ‘annoyed’ and there’s a penalty of $1 per customer.

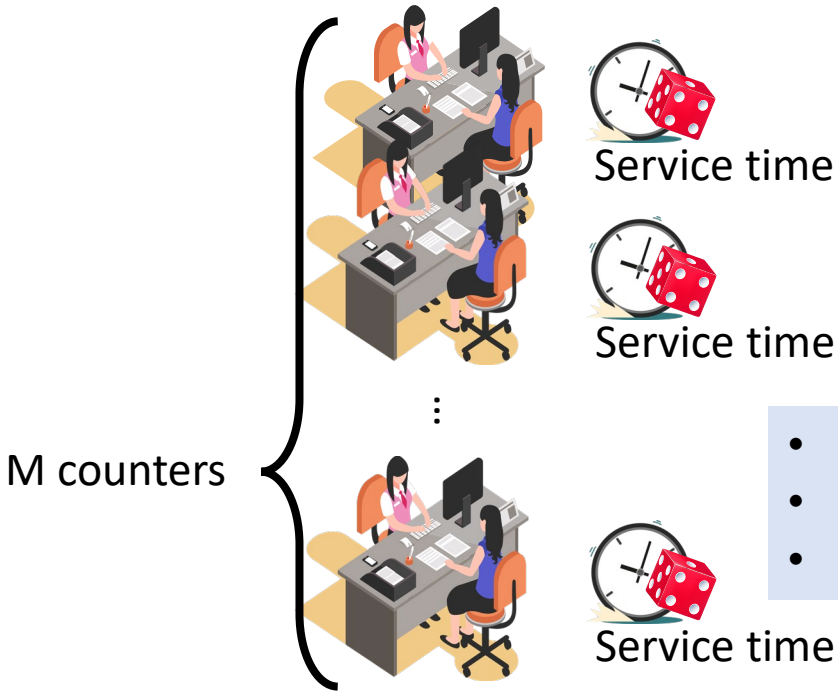
Simulation

Prescriptive Analysis

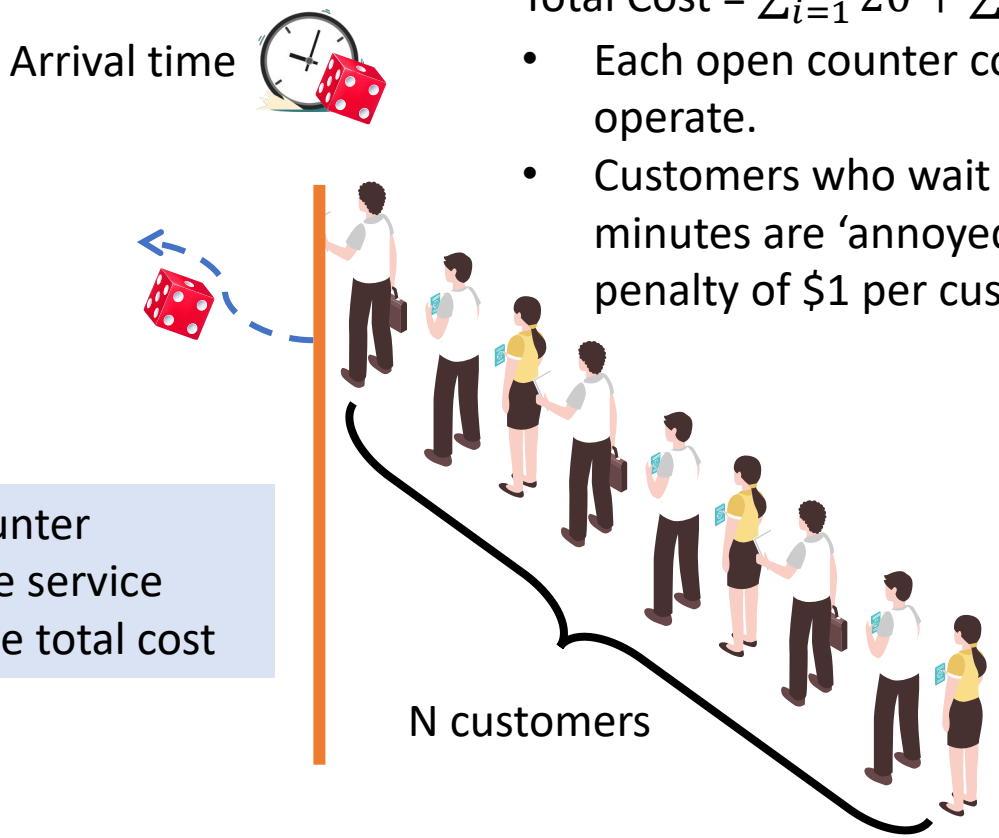
Applying Simulation to Prescriptive Analysis

Example: Optimum number of open Check-out counters

Queue Simulation



- Open M counter
- Simulate the service
- Calculate the total cost



Total Cost = $\sum_{i=1}^{M} 20 + \sum_{j=1}^{N} (t_j^{wait} > 10)$

- Each open counter costs $20 to operate.
- Customers who wait more than 10 minutes are ‘annoyed’ and there’s a penalty of $1 per customer.

Simulation

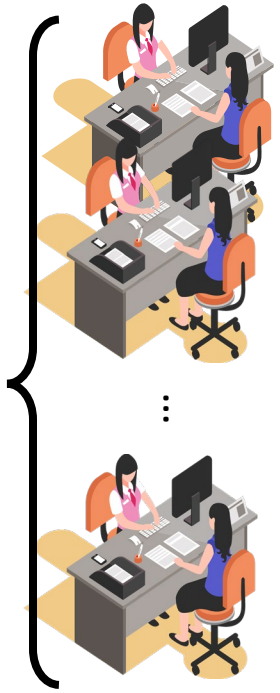
Prescriptive Analysis

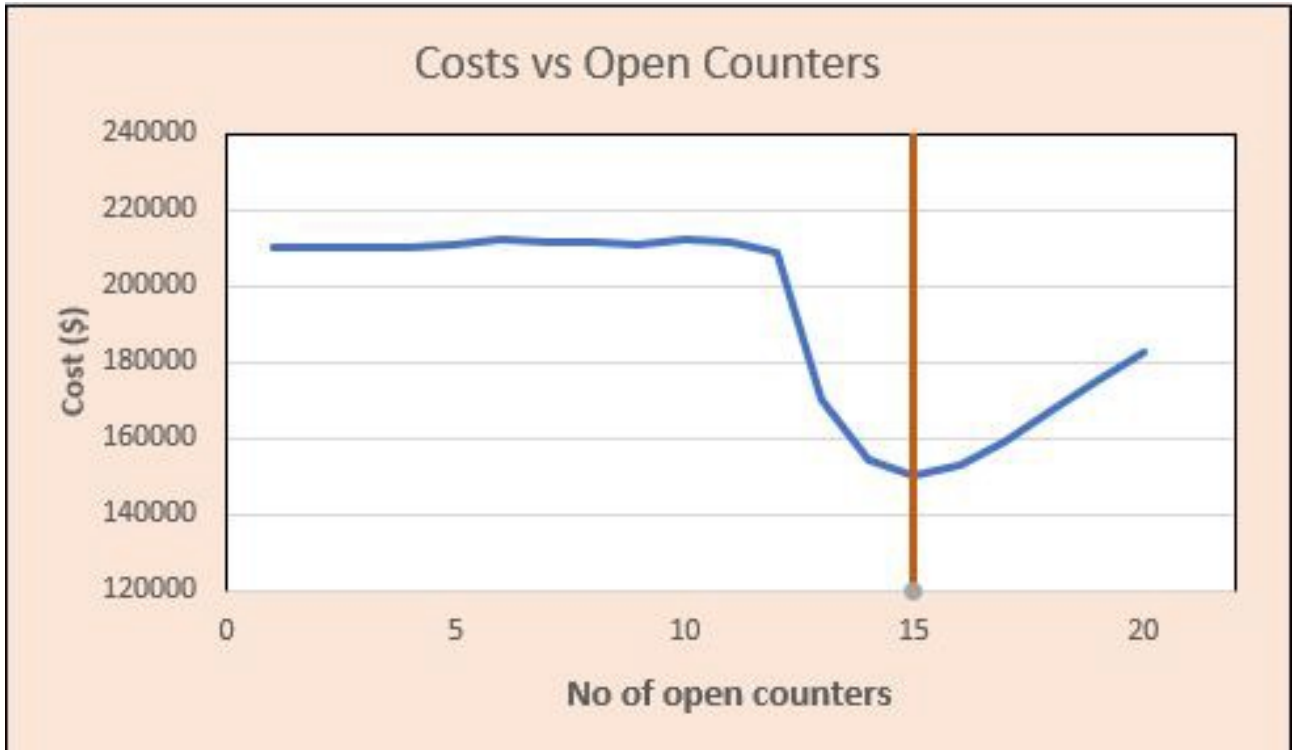
Applying Simulation to Prescriptive Analysis

Example: Optimum number of open Check-out counters

Queue Simulation

15 counters achieve the minimum cost





Source: https://towardsdatascience.com/every-data-scientist-needs-to-read-these-simulation-stories-7be0531e782f

Simulation

Prescriptive Analysis

As can be seen,

Random Generator is invoked in simulation systems.

For example:

Arrival time of each customer

Service time for a customer

} Random under Gamma distributions

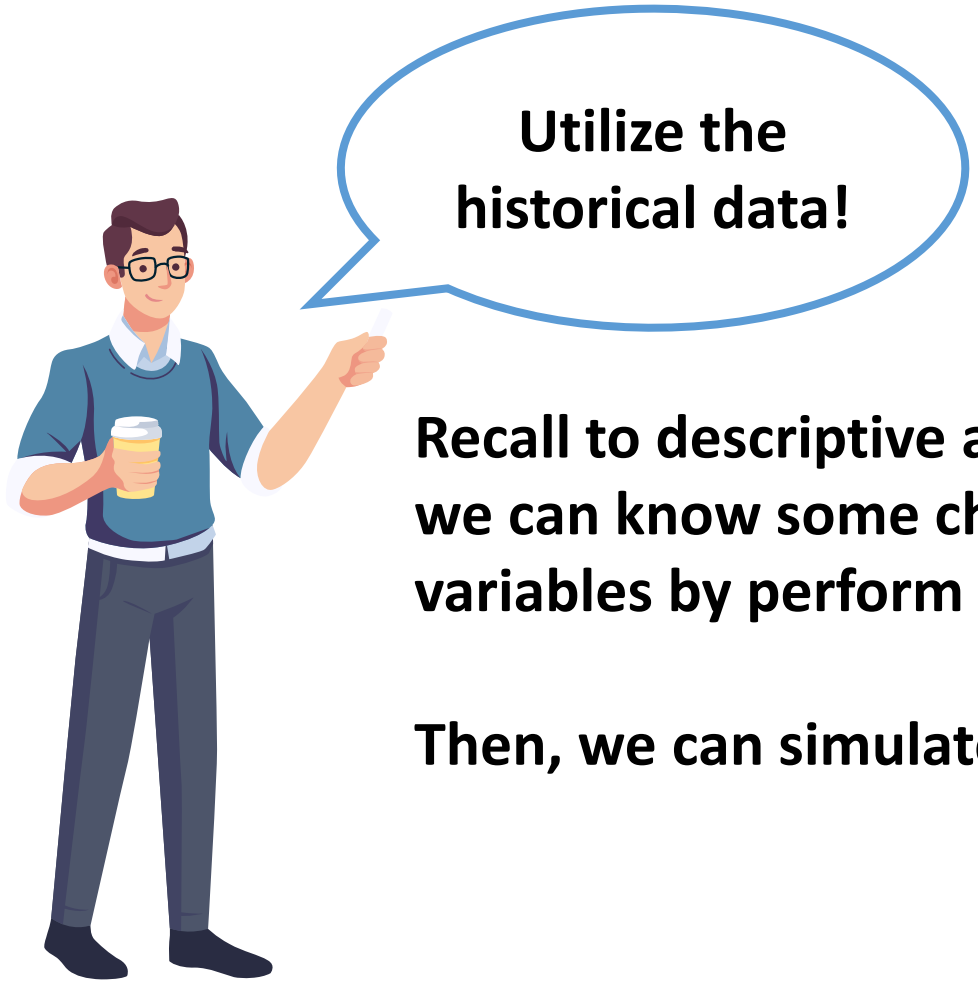
Counter a customer picks → Random under uniform distribution

Expert can tell us what the simulation should be.

If I don't have any expert for the system that I want to construct.

Simulation

Prescriptive Analysis



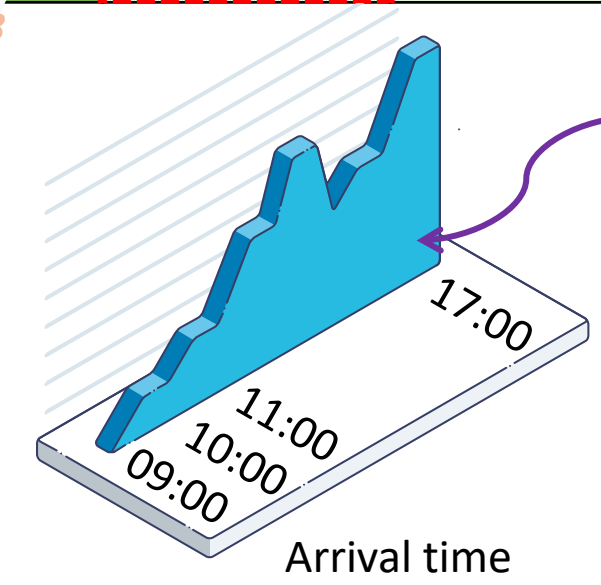
Recall to descriptive analysis,
we can know some characteristics/behavior of random variables by perform descriptive analysis on historical data.

Then, we can simulate a system under the behavior.

Simulation

Prescriptive Analysis

	Arrival time	Service time	End service time	Counter No.
$\mathbf{x}_1$	9:30	9:05	9:20	1
$\mathbf{x}_2$	9:33	9:21	9:30	1
...				
$\mathbf{x}_n$	16:35	16:45	17:00	4



We can generate random numbers under the distribution.

Simulation

Prescriptive Analysis

Monte Carlo simulation

Computerized mathematical technique that allows people to account for risk in quantitative analysis and decision making.

Monte Carlo Casino, Monaco.

Simulation

Prescriptive Analysis

Pattern of Monte Carlo simulation process

1. Define a domain of possible inputs
2. Generate inputs randomly from a probability distribution over the domain

3. Perform a deterministic computation on the inputs
4. Aggregate the results

Simulation

Prescriptive Analysis

Example: Approximating the value of π

1	2	3	4
Define a domain of possible inputs	Generate inputs randomly from a probability distribution over the domain	Perform a deterministic computation on the inputs	Aggregate the results
(x,y) - The values of x and y are in [-1,1]	Random values of x and y are in [0,1] under uniform distribution.	Count the number of points inside the circle.	Approximate the value of π by: $\pi \approx \frac{4 \times N_{inner}}{N_{tatal}}$

See animation at https://academo.org/demos/estimating-pi-monte-carlo/

Simulation

Prescriptive Analysis

Applying Monte Carlo simulation to decision making

Example: Monty Hall Problem

Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats.

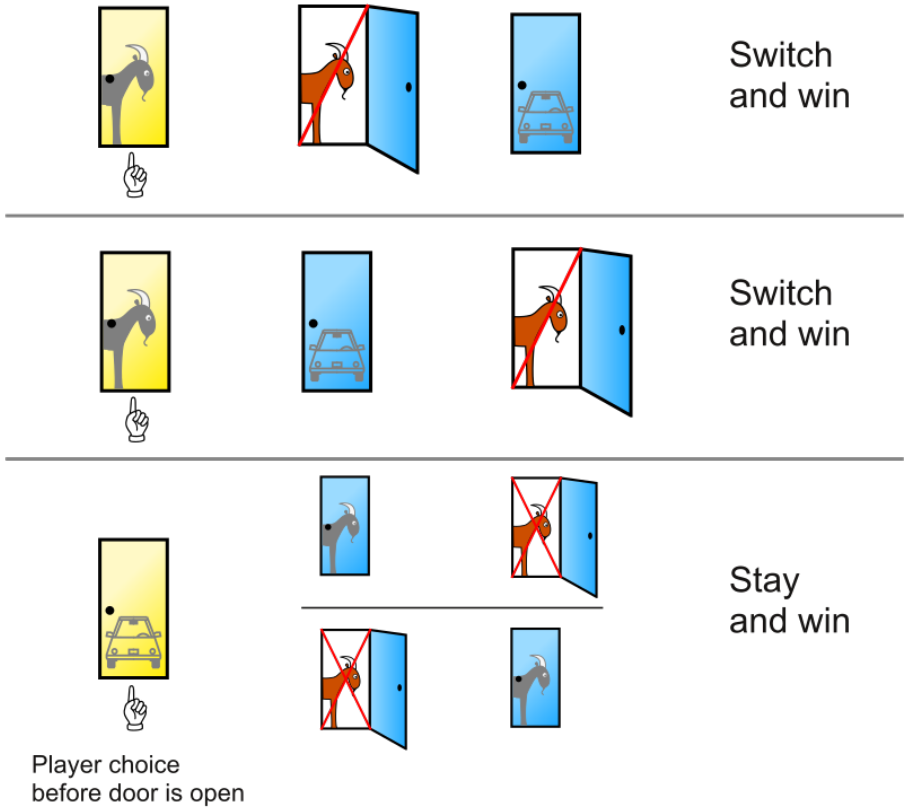
- You pick a door, say No. 1,
- The host, who knows what's behind the doors, opens another door, say No. 3, which has a goat.
- He then says to you, "Do you want to pick door No. 2?" Is it to your advantage to switch your choice?

Simulation

Prescriptive Analysis

Applying Monte Carlo simulation to decision making

Example: Monty Hall Problem



Three initial configurations of the game.
In two of them, the player wins by switching away from the choice made before a door was opened.

Source
https://en.wikipedia.org/wiki/Monty_Hall_problem#/media/File:Monty_Hall_Problem_-_Standard_probabilities.svg

Simulation

Prescriptive Analysis

Applying Monte Carlo simulation to decision making

Simulating Monty Hall with Monte Carlo

1	2	3	4
Define a domain of possible inputs	Generate inputs randomly from a probability distribution over the domain	Perform a deterministic computation on the inputs	Aggregate the results
***N* Doors** – each door has either car or goat, w.r.t. only one door has car. **Initial door** – select 1 from N doors. **Decision** – stay or switch.	Random to put car behind a door. Random the initial door. Random your decision.	Perform *M* times Count the number of success on switch, S_{switch}. Count the number of success on non-switch, S_{stay}.	Calculate Prob. of success if we do not switch doors. Calculate Prob. of success if we switch doors.

Simulation

Prescriptive Analysis

Applying Monte Carlo simulation to decision making

Simulating Monty Hall with Monte Carlo

After we simulate 10000 games of Monty Hall with 3 doors

We got the results as follows:
Number of success on switch: 666,566 - 66.6566%
Number of success on non-switch: 333,434 - 33.3434%

We can see that the success percentages don't vary much, and it tells us that if we make the switch then the chances of winning are 2 out of 3 times.

Further Study

Papers:

- Katerina Lepenioti, Alexandros Bousdekis, Dimitris Apostolou, Gregoris Mentzas, Prescriptive analytics: Literature review and research challenges, International Journal of Information Management, Volume 50, 2020, Pages 57-70, https://doi.org/10.1016/j.ijinfomgt.2019.04.003.
- Lepenioti K., Bousdekis A., Apostolou D., Mentzas G. (2019) Prescriptive Analytics: A Survey of Approaches and Methods. In: Abramowicz W., Paschke A. (eds) Business Information Systems Workshops. BIS 2018. Lecture Notes in Business Information Processing, vol 339. Springer, Cham

Website

- https://towardsdatascience.com/every-data-scientist-needs-to-read-these-simulation-stories-7be0531e782f
- https://towardsdatascience.com/a-zero-math-introduction-to-markov-chain-monte-carlo-methods-dcba889e0c50