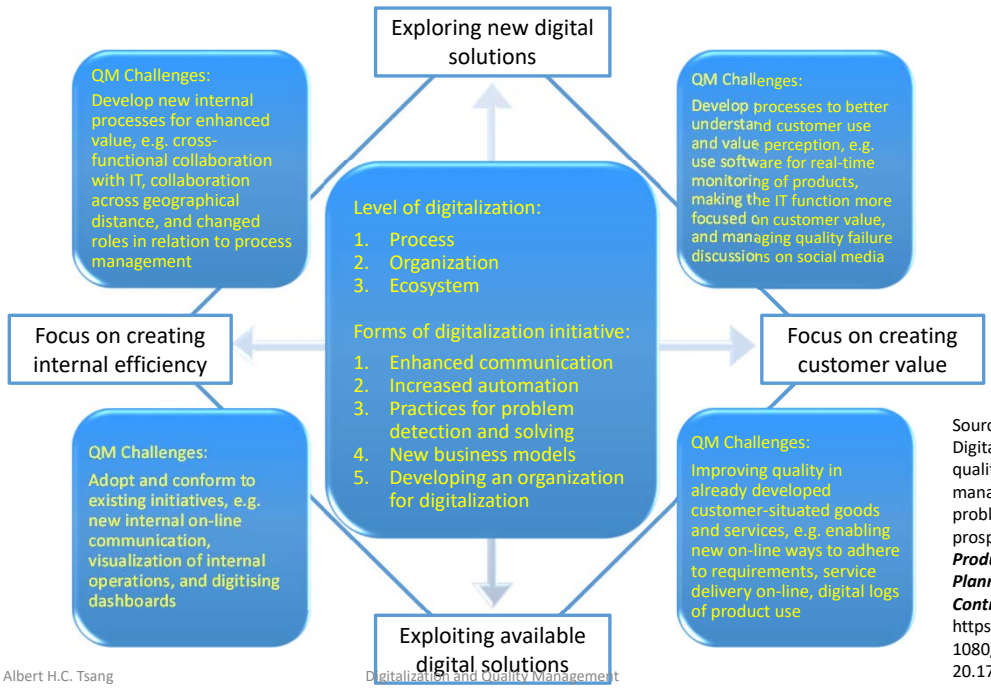
Digitalization and Quality Management

Digitalization

- **Digitization** takes an existing practice or process and move it into a digital environment without changing its content or the actors involved
- **Digitalization** (aka **digital transformation**) involves using digital technologies to change business model, for example, by changing actors in the value chain creation process, create new offerings, or make adaptations to processes , organizations, or ecosystems

Enabling Technologies of Digitalization

- **Internet of things (IoT)**
- **Connectivity**
- **Cloud Computing**
- **Big Data**
- **Data Science**
  - **Analytics Capabilities**
  - **Data Visualization**



Source: Digitalisation and quality management: problems and prospects, ***Production Planning & Control***, 2020, https://doi.org/10.1080/09537287.2020.1780509

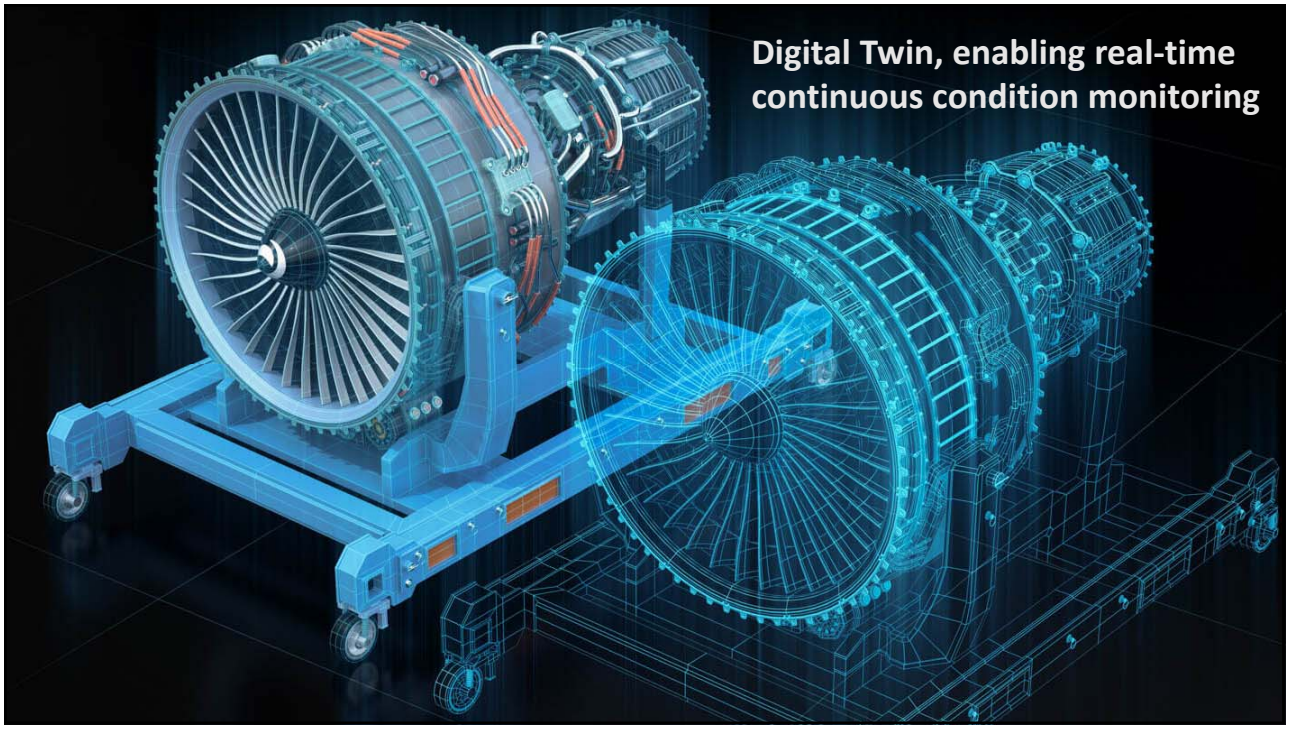
Digital Dashboard

- Aggregates and visualizes data from multiple sources in real time

Data to be added to the dashboard must be:

- Accurate
- Include the KPIs
- Updated frequently
- Consistent
- Traceable

Consider to integrate the dashboard into apps



From Business Problems to Data Mining Tasks

- ***Classification* and *class probability estimation*** attempt to predict, for each individual in a population, which of a (small) set of classes this individual belongs to. An example classification question would be:

  "Among all the customers of TelecoCorp, which are likely to respond to a given offer?" In this example the two classes could be called 'will respond' and 'will not respond'.

- ***Regression*** attempts to estimate or predict, for each individual, the numerical value of some variable for that individual. An example regression question would be:

  "How much will a given customer use the service?"

From Business Problems to Data Mining Tasks

- ***Similarity matching*** attempts to *identify* similar individuals based on data known about them. It is the basis for one of the most popular methods for making **product recommendations**.
- ***Clustering*** attempts to *group* individuals in a population together by their similarity, but not driven by any specific purpose. It is useful in preliminary domain exploration. An example clustering question would be:

  "Do our customers form natural groups or segments?"

From Business Problems to Data Mining Tasks

- ***Co-occurrence grouping*** (aka *market-basket analysis*) attempts to find *associations* between entities based on transactions involving them. An example co-occurrence question would be:
  "What items are commonly purchased together?"
- ***Profiling*** (aka *behaviour description*) attempts to characterize the typical behaviour of an individual, group, or population. An example profiling question would be:
  "What is the typical smartphone usage of this customer segment?"

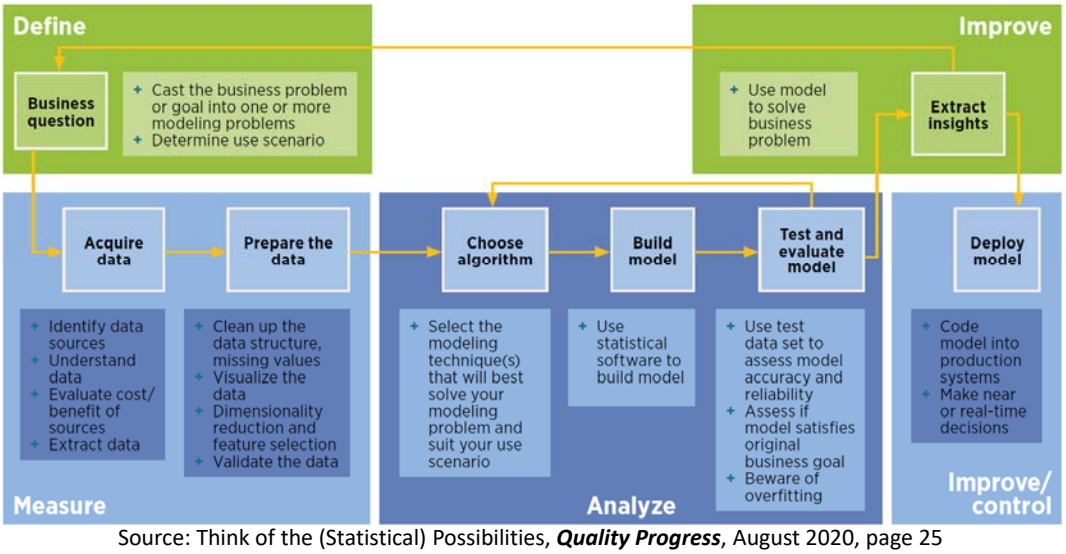
From Business Problems to Data Mining Tasks

- ***Link prediction*** attempts to predict connections between data items, and it may also estimating the strength of the link. Link prediction is common in social networking systems. Here is an example:
  "Since the videos you watched and a new video share the same set of genres, maybe you'd like to watch the new video too."
- ***Data reduction*** attempts to take a large set of data and replace it with a smaller set of data that contains much of the important information in the larger set.

From Business Problems to Data Mining Tasks

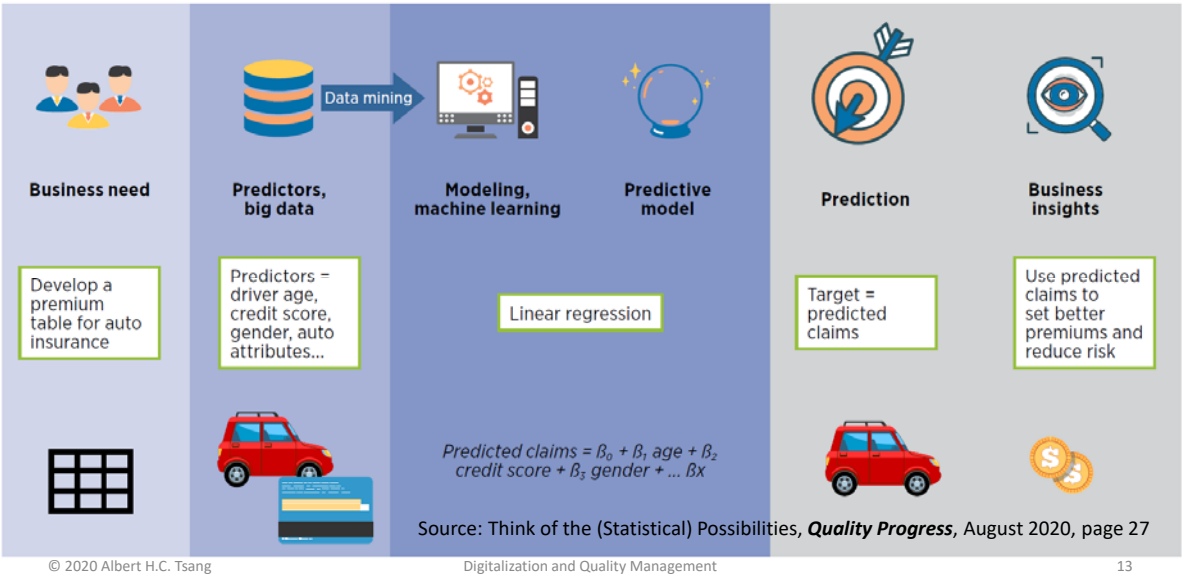
- ***Causal modelling*** attempts to help us understand what events or actions actually *influence* others. It involves the use of randomized controlled tests (RCT) – the so-called 'A/B tests'. The placebo effect in medicine should be designed out in RCTs

Big Data Analytics Process



Source: Think of the (Statistical) Possibilities, ***Quality Progress***, August 2020, page 25

A Financial Services Example



Source: Think of the (Statistical) Possibilities, ***Quality Progress***, August 2020, page 27

Predictive Analytic Algorithms

Supervised Tools (you have an output value you attempt to predict)

- Regression Models (with a numerical target)
  - How much will a customer use a particular service?
- Classifiers (with a categorical target)
  - Which service package will a customer likely purchase if given incentive I?

Unsupervised Tools (you do not have a specific output value)

- Clustering – what items are commonly purchased together?
- Anomaly Identification – profiling to detect anomalies; fraud detection
- Data Reduction Methods – to replace a large data set with a small set of data that contains much of importance information in the larger set

Quality Professionals' Roles in Data Science Projects to the Business:

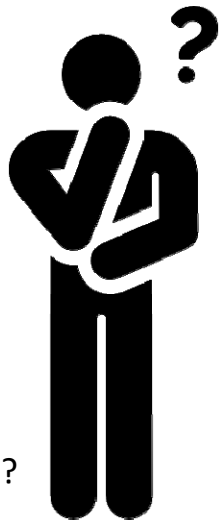
- Connect the work of data science projects to the business
- Understand the business goal behind the project, anchor the project team's work in the context of the broader organizational strategy
  - Data science projects often starts with a question from someone outside the team
  - Discuss and fine-tune questions from stakeholders to better understand the information they actually want and how it will be used
  - Stakeholders might not have a clear idea of what a finished data science product would look like (or how they would apply it)
  - To fill the gap, make sure members of the data science team are regularly invited to production strategy meetings

Quality professionals can be major contributors to development and implementation of data-driven decision making

Quality professionals are well versed in:

- Managing systematic and repeatable processes
- Managing learning and continuous improvement
- Data science and predictive analytics

Answering Business Questions



- Who are the most profitable customers? – use Pareto analysis to rank customers according to profitability
- Is there really a difference between the profitable customers and average customers? – test of hypothesis and confidence interval estimation
- Can we characterise the profitable customers? – use data mining techniques to find patterns automatically
- Will some particular new customer be profitable, how much revenue should I expect this customer to generate?

Predictive analytic algorithms can eliminate cognitive biases in human decision-making:

- Saliency Bias – overly influenced by an analogy to a memorable success
- Loss aversion – to be more sensitive to potential losses than potential gains
- Overconfidence – a manager over committed and spent significantly more resources on a pet project

The algorithm itself does not make biased decisions.
It reproduces biases that were built in the training data.

Reinforcing War Planes

- In WWII, the vulnerability of airplanes to enemy fire was examined
- Some parts of airplanes were hit disproportionally more than other parts
- Conclusion: these parts should be reinforced
- Right or wrong?
- Wrong! The parts that are hit least often should be reinforced

An example of Selection Bias

A plane would be less likely to return if it were hit in a critical area

Boston's Street Bump Project

- A project of Boston's Mayor's Office:

  An app called *Street Bump* was designed to collect smartphone data from drivers – a low cost way to collect information on pothole repair
- A lot of false positives were received because the app has difficulty distinguishing between bumps in the road, manholes, and potholes
- A worse problem:

  By relying only on feedback from the app, Boston was not receiving information from neighborhoods where residents didn't own smartphones

Propensity to Heart Diseases

- African Americans are more likely than white Americans to have heart diseases
- Black people are more likely to have heart diseases, right?
- However, black people in Africa are not significantly different from whites in the likelihood of having heart diseases
- A painting about slave trader in the colonial era provides a clue
- Slaves with saltier sweat were more likely to survive the grueling sea journey across the Atlantic
- However, these people were also more likely to have hypertension, a hereditary health problem

An Anomaly in American Poultry Farms

- 100% of the turkeys bred in US poultry farms are conceived by artificial insemination, why?
- Interestingly, a significantly percentage of chickens bred in US poultry farms are conceived naturally, why?
- The anomaly is due to the over-sized breasts of these turkeys that are preferred by US consumers; making it impossible for them to have sexual intercourse
- Lesson learnt – data alone may explain What has happened, but cannot tell us Why
- To know why, we need to go to the Gemba – the crime scene

About the Speaker

albert.tsang@polyu.edu.hk

Albert H.C. Tsang is an advocate of quality and reliability in Hong Kong. He is the representative of American Society for Quality (ASQ) in Hong Kong, a former Chairman, founding member, and Fellow of Hong Kong Society for Quality (HKSQ). He had developed and conducted many customized training courses on various aspects of quality and engineering asset management for many organizations and professional bodies in Hong Kong, the Americas, Middle East and South Africa. He has also provided consultancy services to organizations in the public, governmental, business and industrial sectors on matters related to quality, reliability, maintenance and performance management.

Dr. Tsang is a co-author of the best-selling book: ***Maintenance, Replacement, and Reliability: Theory and Applications***, the 2nd edition of which was published in June 2013. He is also the author of "***WeibullSoft***", a computer-aided self learning package on Weibull analysis.

