

BI – Preliminary Information on the course

Prof. Paola Velardi

Dipartimento di Informatica

Sapienza Università di Roma

He course has 2 parts

- Business Intelligence and Social Networks 6 CFU, Paola Velardi
- Business Process Modelling, 3 CFU, Paolo Bottoni

- BI from September 18th to November 15°
- Written test (only for those who attend in presence) on November 21st
- The course has a web site where you find everything, READ CAREFULLY

<https://twiki.di.uniroma1.it/twiki/view/BI/WebHome>



[WebHome](#) < [BI](#) < [TWiki](#)

Overview

Presentation of the course
[CourseIntroductionandObjectives.pdf](#)

Enterprises today are driven by data.
"Business Intelligence allows people of all levels in organizations to access, interact with, and analyze data to manage the business, improve the performance, discover opportunities, and operate efficiently" (Cindi

Schedule and Exams (read carefu

Class days (2022-23)

To be defined

Class days (2021-22):

Classes will start on Monday, September 27th, 2021

It is highly recommended to attend the

Suggested Text Books

PART A

- Cindi Howson "Successful Business Intelligence" Second Edition, Mc Graw Hill
- Ramesh Sharda, Dursun Deelen and Efraim Turban "Business Intelligence: A Managerial Perspective on Analytics" Third Edition, Pearson
- Rick Sherman "Business Intelligence

Part A: Course Material

IMPORTANT NOTES:

- **Subscribe** to Google Group (Prof. Velardi) [Business Intelligence Google Group 2022-2023](#)
- **LABS:** Labs on Watson Studio will be held starting on October. Details in class.

Part B: Process Modeling

IMPORTANT NOTES:

Subscribe to the Google Group of BI / Process Modeling (Prof Bottoni) using your institutional address (domain: [@studenti.uniroma1.it](#)): <https://groups.google.com/g/bpmbottoni>. Teaching material will be posted on a [dedicated remote Google Drive folder](#). The

Use cases, datasets and readings

Open Data sources for Business Intelligence and Business Process Intelligence

- [Data resources for BI: a very large catalogue of freely available datasets](#)
- [Another pointer to open data sources](#) for business intelligence
- The Business Process Intelligence

Why this course

- Enterprises today are driven by data. "*Business Intelligence allows people of all levels in organizations to access, interact with, and analyze data to manage the business, improve the performance, discover opportunities, and operate efficiently*" (Cindi Howson, Successful BI, [McGrawHill](#)).
- However, the degree to which BI solutions can be successfully adopted within organizations depend to a great extent on the degree to which **business and IT experts can partner together**.
- **The objective of this course is to form Advanced Business Users of BI applications**, with a deep understanding of the business needs and some (even practical) understanding of technology.
- The *Advanced Business User* **understands the business and how to leverage technology to improve it**, leads the interpretation of business requirements and strategic objectives, and helps designing reports **to answer business questions**.

Organization

- The course is in two parts:
 - PART A: Business Intelligence and Social Analytics (Instructor: Prof. Paola Velardi velardi@di.uniroma1.it) 6CFU (MON-TWE-WED from september to mid november)
 - PART B: Business Process modeling (Instructor: Prof. Bottoni) 3CFU (same days, mid november to end of semester)
- Material and notes sent via email and shared DRIVE folder: [Subscribe to Google group!](#)

Part A (BI)

- About 2/3rd theoretical lessons, 1/3rd laboratory (7 labs)
- Written test last lesson **mid-november - likely, the 21st**- (if not passed, then on January/February 2 regular exam sessions)
- Student seminars and homeworks → **up to +2** on **final** grade (BI)
- Project on Watson Studio (IBM platform for Business Intelligence). **Important to attend labs !** you can start working on your project under the assistance of IBM tutors
- Project must be delivered **WITHIN September 24 session** – but can't register your grade until project is delivered. **Delivery deadlines** will be communicated.
- Final grade BI: test (60%) project (40%). Upgrades depending on seminars. Final grade **total**: 2/3 BI 1/3 BPM. BPM has only a written project, no test.
- *If you attend the course systematically, you get everything done by mid november (my part, BI).*

Theoretical lessons

Topics are presented in class



At home, you study business cases and work on self-assessment questions (sent via google group)

Labs

- Held in class at LabInfo (better to **BRING YOUR PC!!!!**)
- 1 IBM tutor (dr. ing. Flavio Aicardi) + me
- Using cloud version of IBM Watson Studio (6 months free license, agreement Sapienza-IBM)
- Objective: how to apply Watson Studio on everyday business
- Work on **real datasets, real problems**
- You then find your own dataset/business questions and present a project



Bring you PC possibly each
lesson

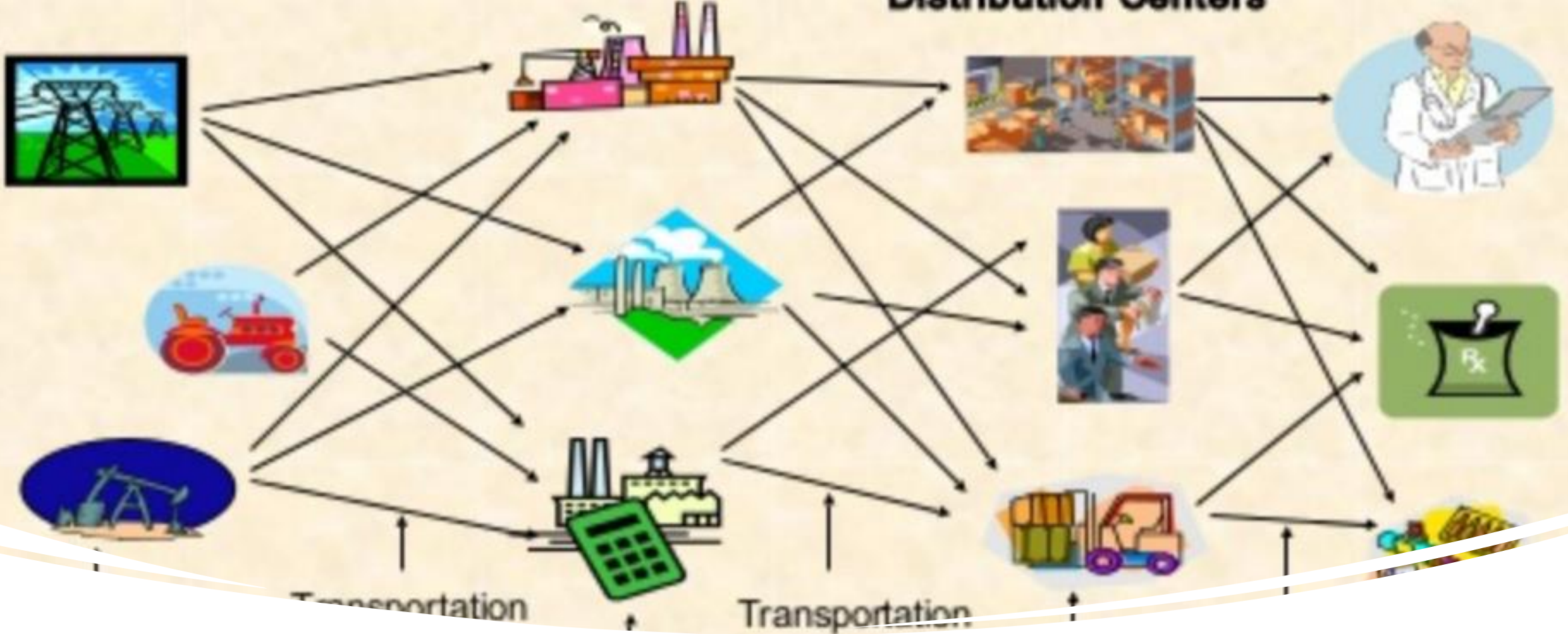
We can work in class!!



Business Intelligence

What is and what is not





Example: supply chain management

- *Supply chain management* is the integration of the activities that procure materials and services and deliver them through a distribution system

Example: supply chain management

- It typically implies the following actions/actors:
 - **Transportation vendors**
 - **Credit and cash transfers**
 - **Suppliers & Distributors**
 - **Accounts payable and receivable**
 - **Inventory**
 - **Order fulfillment**



The traditional Supply Chain workflow

The traditional supply chain...



The
traditional
ICT tools for
handling
SCM

ON-LINE TRANSACTION PROCESSING or
OPERATIONAL TRANSACTION PROCESSING

- ❑ TERMINOLOGY: “**On line transaction processing**, or **OLTP**, is a class of [information systems](#) that facilitate and manage transaction-oriented applications, typically for data entry and retrieval [transaction processing](#). “ (wikipedia)
- ❑ Example: billing and invoice management systems, on-line reservation systems..
- ❑ OLPT workloads are characterized by *small, interactive transactions* that generally require sub-second response times.

Example of a typical OLTP interface

Shipment number, status, carrier name, shipping state, date, source, destination, package type..

Shipment Number	Is Master	Status	Carrier Service Name	Ship To State	Ship Date
MSTR182	✓	Completed	RLC	UT	2/29/2012 1:
102996		Completed	Yellow Roadway	CA	7/11/2012 1:
DR389		Ready To Ship			3/2/2012 4:5
MSTR186	✓	Completed	CTS	NV	3/6/2012 5:1
DR416		Completed	RLC	UT	3/6/2012 4:4
DR385		Completed	FedEx International Economy Freight		3/2/2012 2:5
DR261		Completed	FedEx 3Day Freight	CA	2/22/2012 4
MSTR193	✓	Completed	CTS	TX	3/13/2012 1:
DR448		Completed	CTS	AZ	3/13/2012 1:
DR417		Completed	CTS	AR	3/6/2012 4:5
DR333		Completed	CTS	AR	2/29/2012 2:
DR415		Completed	CTS	UT	3/6/2012 4:5
DR418		Completed	CTS	TX	3/6/2012 4:5
DR331		Completed	FedEx 2Day Freight	MA	2/29/2012 1:
DR388		Ready To Ship			9/4/2012 12
MSTR194	✓	Completed	Yellow Roadway	CA	7/12/2012 6
DR435		Completed	CTS	CA	3/8/2012 9:0
DR186		Completed	RLC	NV	2/15/2012 1
MSTR191	✓	Completed	Yellow Roadway		3/8/2012 10
DR463		Completed	Yellow Roadway	NB	8/2/2

Shipment Number	Status	Carrier Service Name	Ship To City	Ship To State	Ship Date	Package
102986	Completed		Omaha	NB	7/12/2012	1
102986	Completed		Princeton	NJ	7/12/2012	1

SequenceNumber	PackageType	Weight	UOM	TrackingLink	Print Label	Created By	Created
1	CUSTOM	400.00	Pounds	PRO1596543	Print Label	junk@accellos.com	7/12/2012

Little more than an Excel file..

Shipping Metrics Deliveries Admin

Powered by Accellos One Workspace 2.6.7.4 Accellos, Inc. 2012. All rights reserved.

What kind of info can you get from this database?

- Shipments (transactions)
- Destinations
- Earnings, etc.
- Everything according to time, geographic areas, etc. (e.g., “*How many shipments to Roma in July?*”)
- Data come from **inside** the company and are mostly **manually entered**. Data processing is with traditional database systems (e.g., SQL). Interfaces are sometimes well designed , but usually not so exciting.



**CAN BUSINESS INTELLIGENCE HELP SUPPLY
CHAIN MANAGEMENT?**



Study Case one: Magpie Vaccine supply chain (1)

- **Cold chain** in healthcare is the temperature-controlled supply chain in transporting and storing vaccines and drugs
- Maintaining cold chain is extremely important for these products
- A study of CDC (Center for Disease Control in US) revealed that $\frac{3}{4}$ of providers experienced serious cold chain violation
- This has been a very serious issue during pandemics

Study Case one: Magpie Vaccine supply chain (2)

- **Magpie sensing** (a start-up project) has been one of the first companies using cold chain **data analytics** and **business intelligence** to improve vaccine supply chain management
- They gather location-tagged information about cold storage units using a **shippable wireless temperature and humidity monitor** system
- Real-time, **fixed-location monitoring** devices then safeguard products from their arrival at a lab or clinic until the moment of use





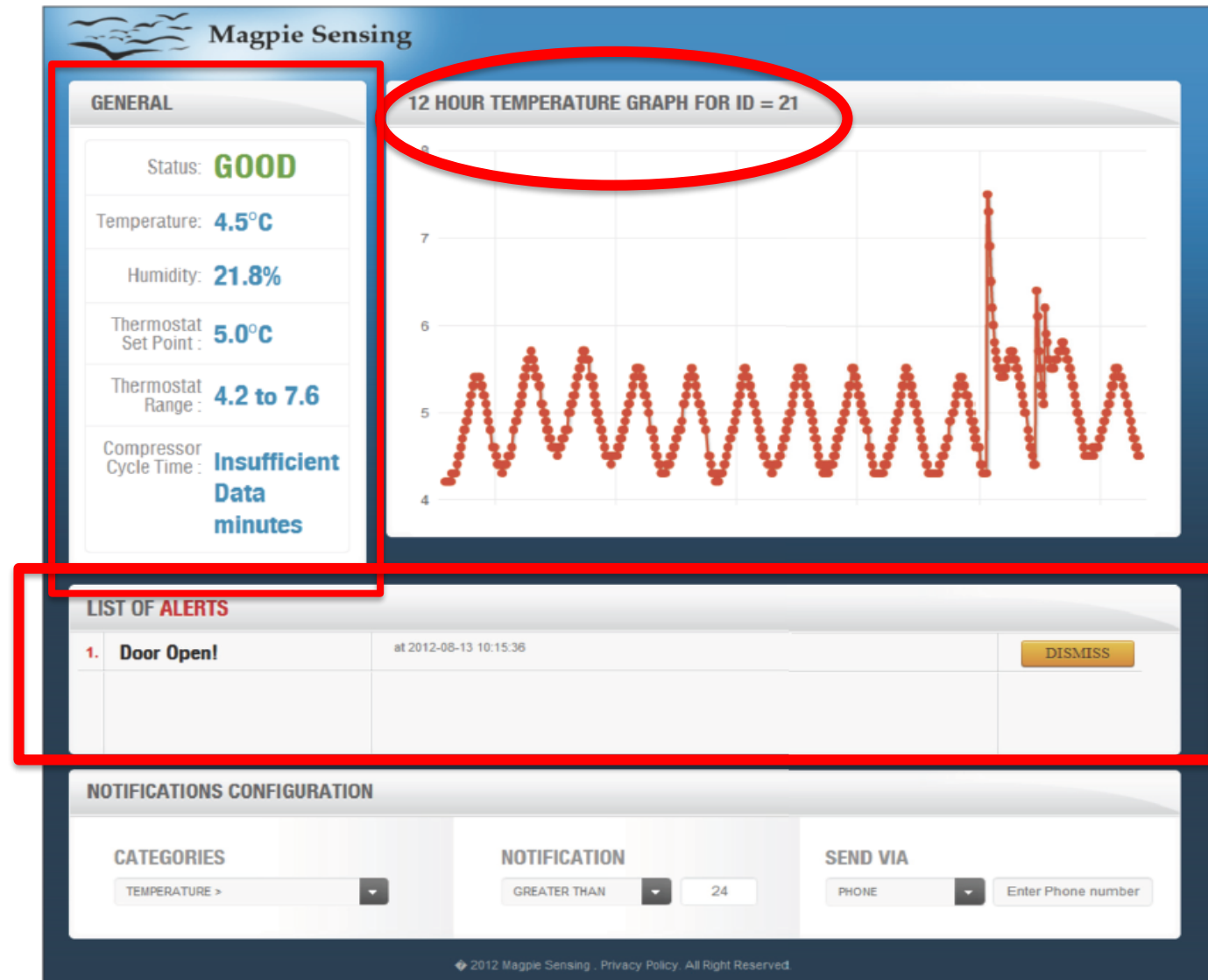
Study Case one: Magpie Vaccine supply chain (3)

- At the core of Magpie Sensing's solutions are data **analytics algorithms** which leverage data from monitoring devices to improve cold chain processes and **predict cold storage problems before they occur.**

Study Case one: Magpie Vaccine supply chain (4)

- Three types of analytics are provided:
- **Descriptive:** set point of the system's thermostat, typical range of temperature values in the system (by season), the duty cycle of the system's compressor, etc. This **informs** the user as to whether their storage unit is properly configured to store a particular product.
- **Predictive.** Predictive analytics **detect** cold chain problems (e.g. poor equipment configuration, human error, equipment failures) and send an alert **before** temperature bounds are violated.
- **Prescriptive:** provide **recommendations** to improve cold storage processes and business decisions during normal operation (e.g., optimal thermostat setting for a particular product)


Example of analytic system: single device



What's new?

- Data come from **several sources** (in addition to company data, we have sensors , methereological and geographical data)
- Data are aggregated to obtain **additional information** (e.g., times of day that are busiest and the periods where the system's door is opened the most. This can guide staff training and institutional policies)
- Not simple queries (descriptive) but also **predictive** and **prescriptive** information. System helps efficient and timely **decision making!**
- Interfaces (**visualization**) is a key to efficient use and understanding of data

Study Case
two:
Hokuriku
Coca-Cola
Bottling
Company (1)

- When you select a beverage on a vending machine, you never think that there might be a lot of hard work and careful planning by the bottling companies, who have to decide which drinks will sell best in their machines while at the same time eliminating out-of-stocks and reducing equipment failures.
 - Hokuriku has used big data and data analytics to improve efficiency of vending machines. They reported to increase 10% sales and decrease 46% associated costs.
- 

Study Case two: Hokuriku Coca-Cola Bottling Company (2)

- The first step was to provide vending machines with **wireless connections** to transmit data
- Data from vending machines and other traditional sources are collected in a **data warehouse** (will see what this is later: for now, TERMINOLOGY: “a large heterogeneous database of structured and unstructured data”)
- Warehouse collects data like: *time location* and *date* of each sale, when a product *sells out*, or is *short-changed*, and when machine is *malfunctioning*. They also collect *sentiment data* from social networks.
- Big data technologies (SAP HANA, Hadoop, and Spark) have been used to manage and intersect these data providing **descriptive, predictive** and **prescriptive** analytics



コカ・コーライーストジャパン
多摩工場
Tama Plant *Coca-Cola East Japan*

Welcome

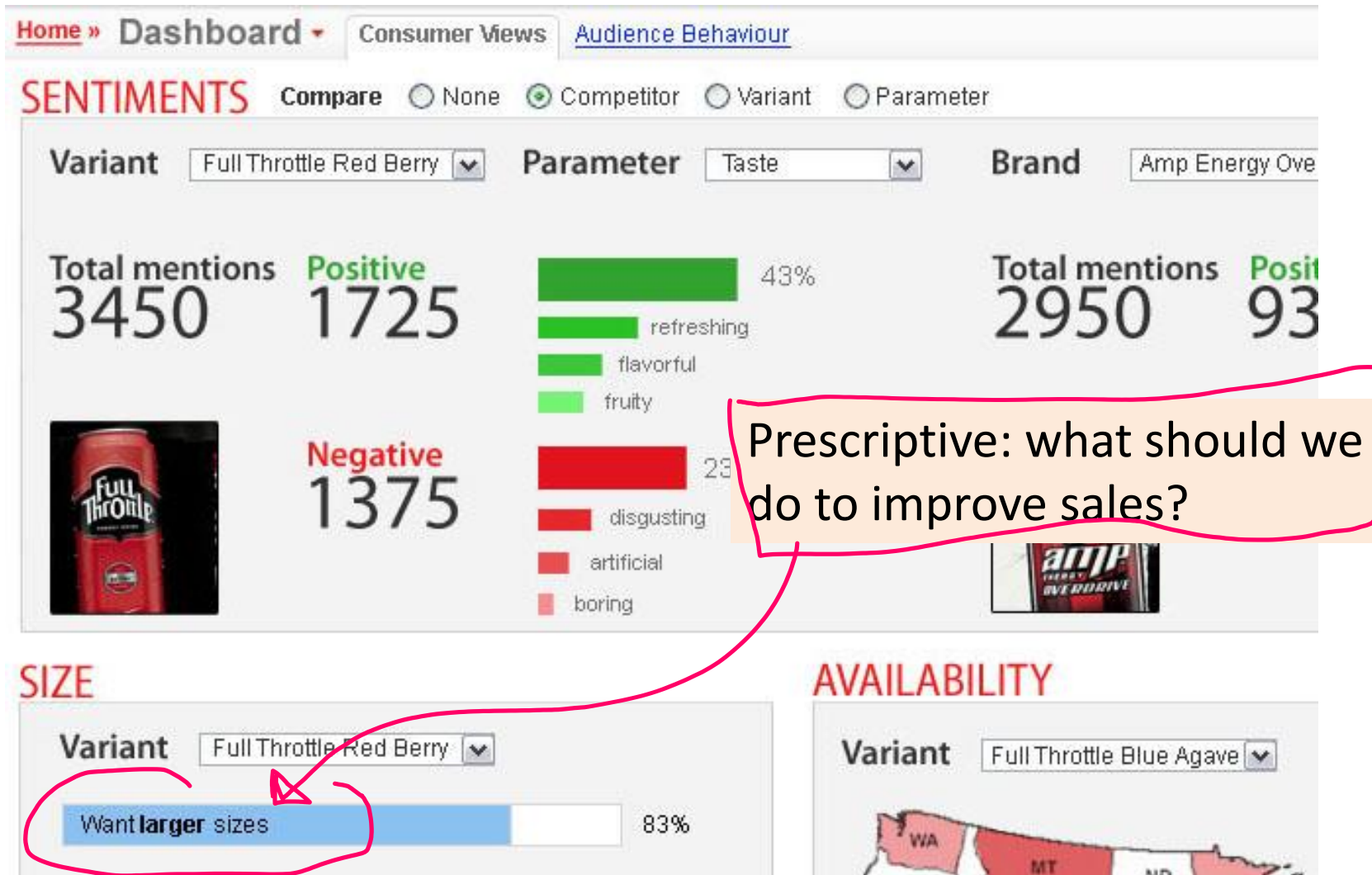


Example of analytic system: brands

- Descriptive: what is the best sold?
- Predictive: which products are close to be out-of stock?
- Prescriptive: how can we sell more of a give beverage?



Example of analytic system: audience behaviour



Quiz time:
Can you
imagine a
PREDICTIVE
question this
system may
answer?

- Remember: «predictive» means a question regarding something that may happen in the future. Knowing in advance the possible occurrence of an event may help your business!



What's new?

- Data come from **several sources** (in addition to company data, we have near real-time data from single vending machines, and social data)
- Data are aggregated to obtain **additional information** (e.g., best seller by time and location, opinions of users, user's preferences by profile, date, location... This can guide stocking policies and marketing campaigns)
- Not simple queries (descriptive) but also **predictive** (a product will shortly sold-out, fine-tuning policies for refilling: not only if close-to-sold-out, but also based on historical data, consumers' opinions..) and **prescriptive** information (targeting users for specific campaigns). System helps efficient and timely **decision making!**
- Interfaces (**visualization**) is a key to efficient use and understanding of data.

Case Study 3

- Lennox International is a \$3B manufacturer and distributor of *heating, air conditioning and refrigeration* products.
- Lennox describes how it **doubled the size** of its footprint with
 - integrated Warehouse Management System (WMS),
 - Transportation Management System (TMS) and
 - Labor Management System (LMS).
- See more at:
<http://www.manh.com/resources/videos/lennox-internationals-supply-chain-integration#sthash.1kjODMBd.dpuf>

Case Study 3: Lennox



HOW LENNOX MASTERED THE SUPPLY CHAIN IN A CHANGING NETWORK

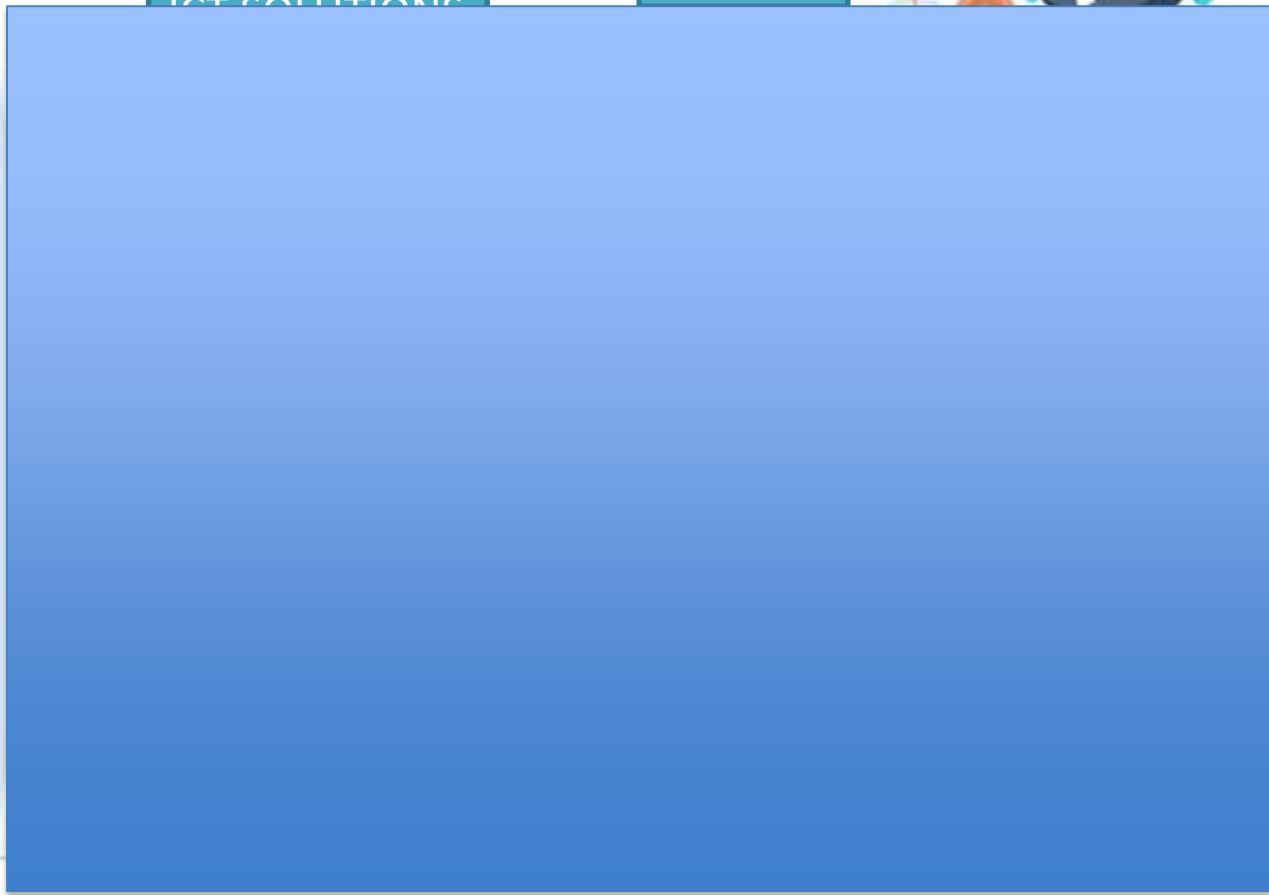


RESULTS

CHALLENGES

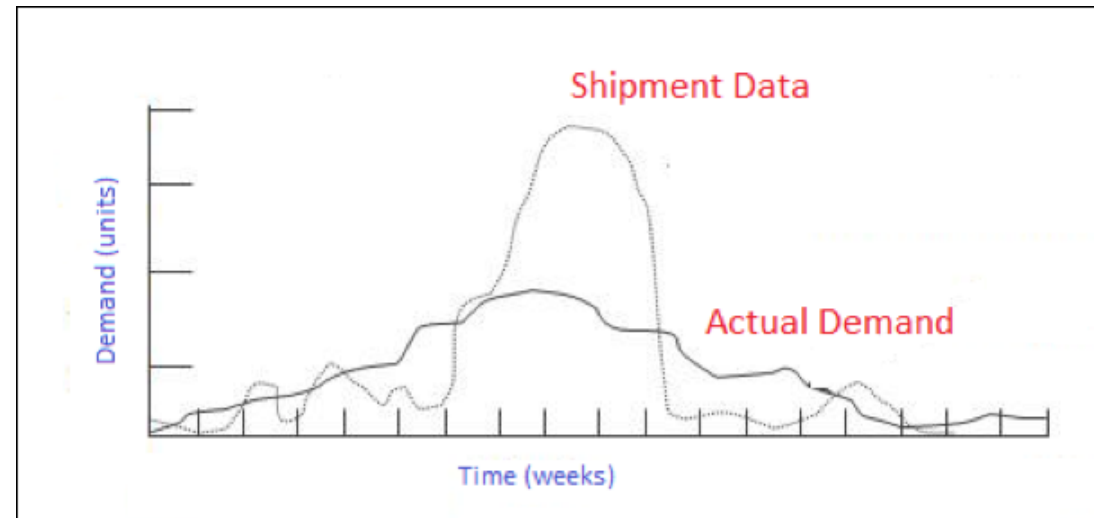
CHALLENGES

- Moving from a national stocking model to a hub-and-spoke model with 55 shipping and 161 selling locations
- A 250% increase in physical locations
- 450,000 SKU-Locations consisting of both parts and finished goods
- High product availability targets- Some customers need 'Same Day' pickup
- Highly variable independent demand
- Highly seasonal demand (AC and heating)
- Many new product introductions



Case Study 3: Lennox (2)

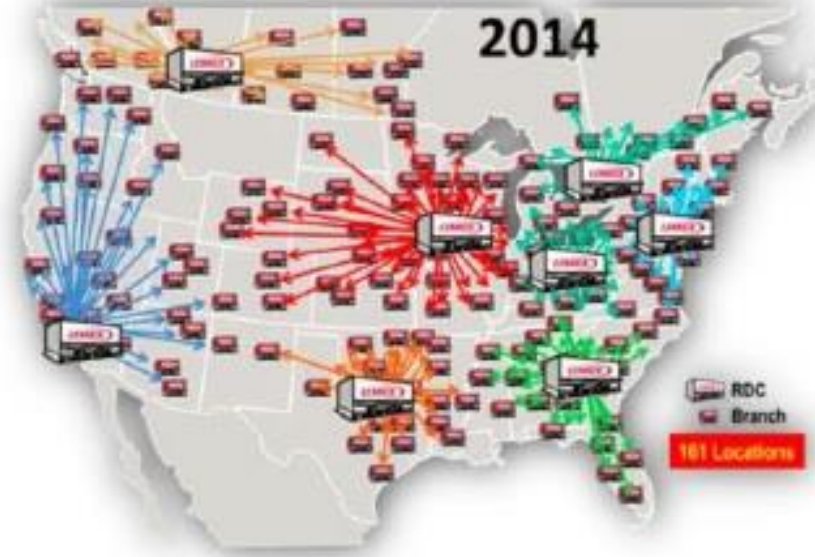
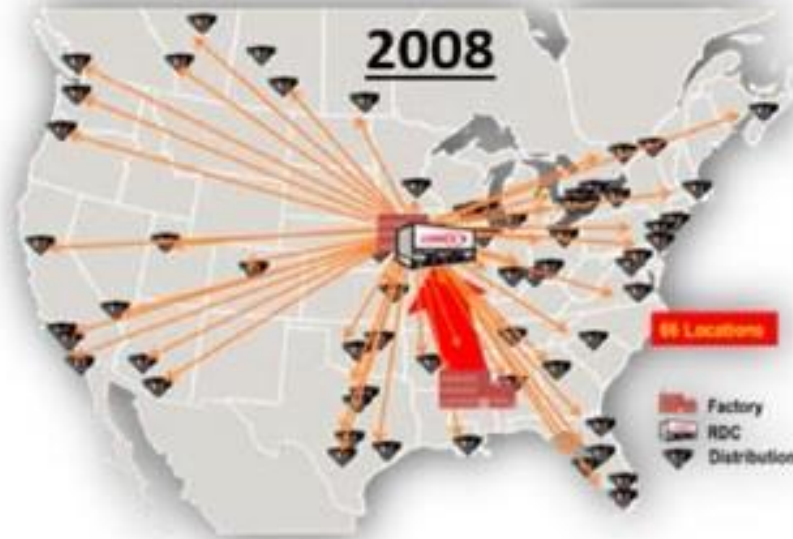
- As for Magpie and Coca Cola, use telemetry data (wireless transmission of data), e.g. for tank volumes.
- Use also POS (Point Of Sale) data rather than shipment data only.
- POS data allows to see the real time data driving the supply chain (what is sold, rather than what is shipped). Statistical methods can then be applied to bin and cluster the data for further analysis.



Case Study 3: Lennox (3)

- **Prescriptive:** Determine which distribution centers should service which stores for which specific products, especially in large networks, in order to *optimize time, fuel and resources*.
- **Prescriptive:** Create the **most productive delivery schedule** for fleets, taking into account factors such as opened and closed stores, route disruptions and seasonal issues
- **Predictive:** Factor in delivery window optimization, predicting the **best time for product delivery** based on store hours, traffic and other variables.
- NOTE: integrating real time data from POS, meteorology, and local sensors is crucial here.

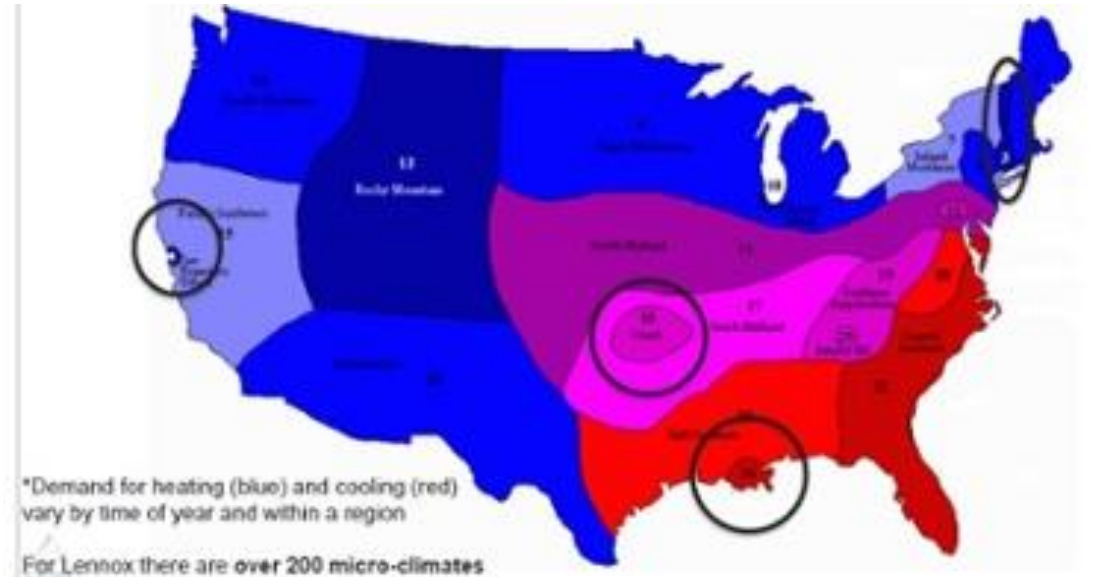
Case Study 3: Lennox (4)



- Lennox almost **tripled the percentage of orders** that can be delivered the next morning, yet reduced inventory by almost 20%, even with a 250% increase in physical locations.
- They improved service levels by **about 20%**. They also reduced distribution costs as a percentage of sales by over 15%, increased inventory turnover by more than 20%, and reduced response time.

Analytic system example: view seasonal demand

Demand for heating (blue) and cooling (red)



(**Descriptive:**) Seasonal demand **at the item level** is identified in specific locations within regions (system identified over 200 micro-climates).

What's new?

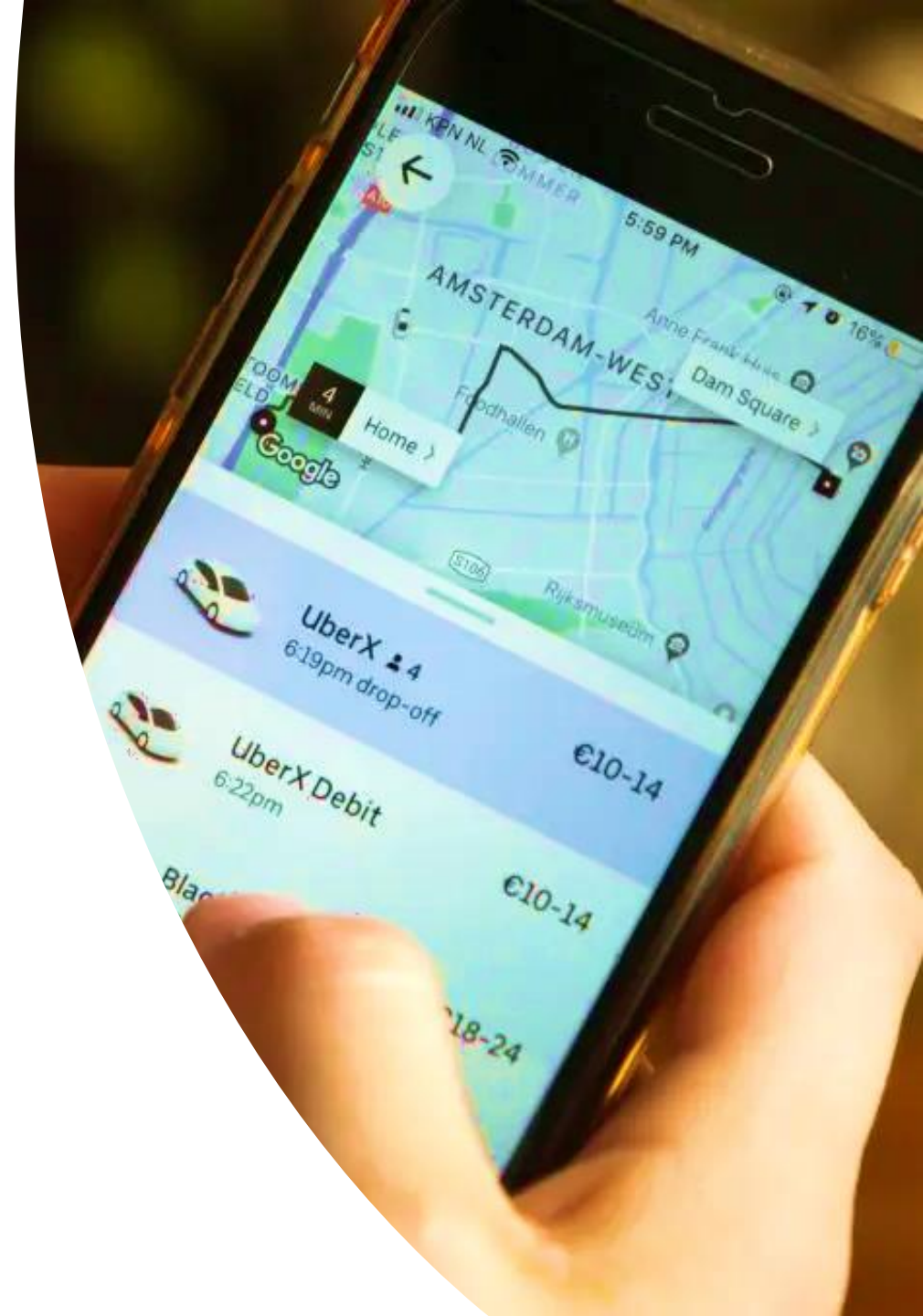
- Data come from **several sources** (in addition to company data, we use POS data and telemetry)
- Data are aggregated to obtain **additional information** (e.g., high-demand products and parts that must be available immediately; lower-demand items need to be delivered the same day or next morning. This can guide staff training and institutional policies, such as accurate stock replenishment at the order-line level)
- Not simple queries (descriptive) but also **predictive** and **prescriptive** information. Advanced analytical techniques to accurately model and forecast complex and unpredictable demand, e.g., according to micro-climates . System helps efficient and timely **decision making!**
- Interfaces (**visualization**) is a key to efficient use and understanding of data

Use case 4: UBER

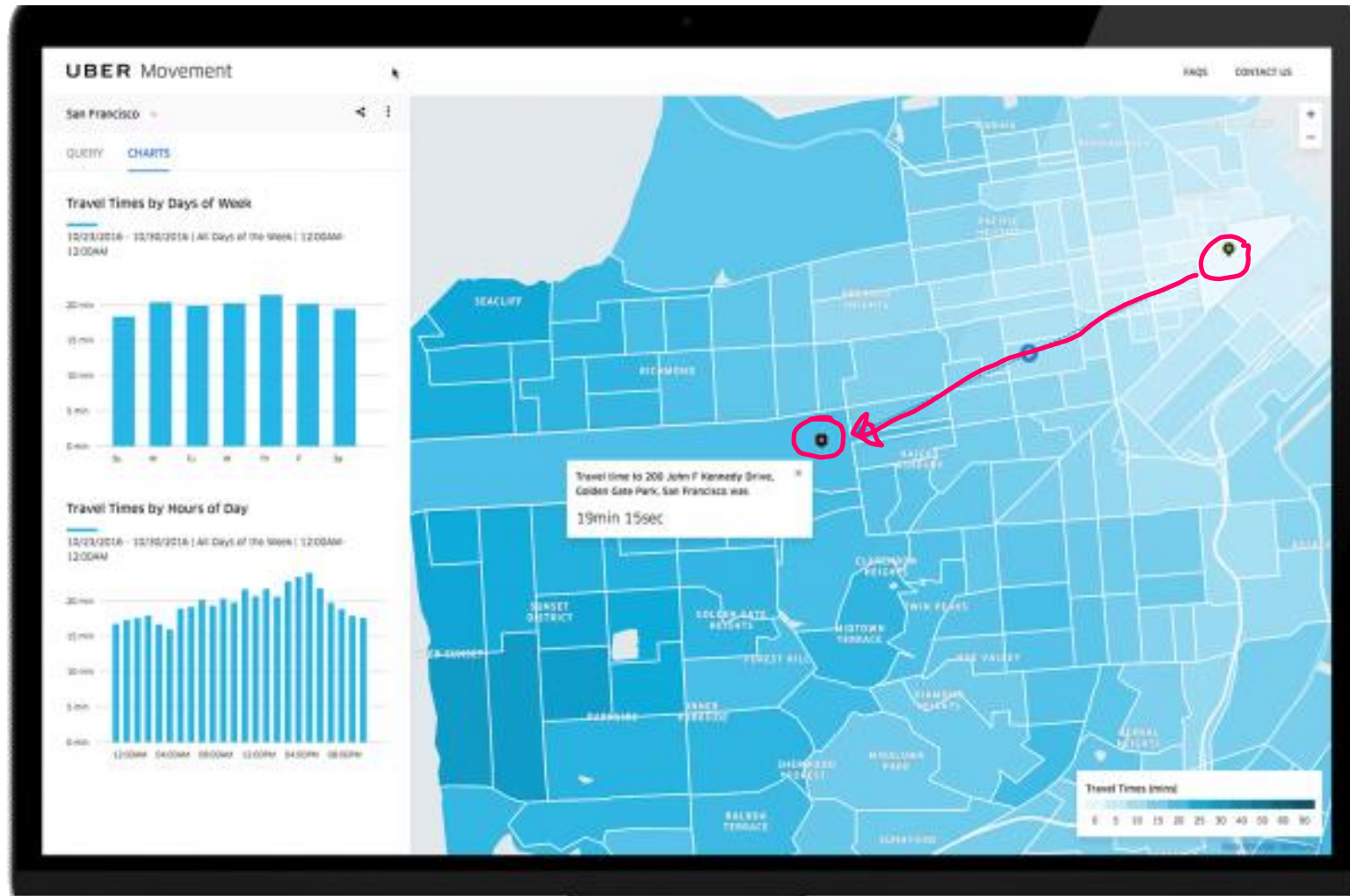
- Which data? reservations from users (through apps), traffic and meteo data, users' ratings of the service/drivers, internal data (staff, revenues,...)

3 types of results:

- Descriptive analytics: revenues, reservations by time, location, day hours, type of transfer medium (car, byke..)
- Predictive analytics: predict supply and demand to adjust for bottlenecks and other common issues.
- Prescriptive analytics:
 - Manpower optimization: let drivers with best users' rating work the most;
 - Price optimization: manage fare prices by increasing the fares when the demand goes up. When there a poor traffic or when the demand for taxis goes up, the prices rise up while the prices reduce significantly when the demand goes down.

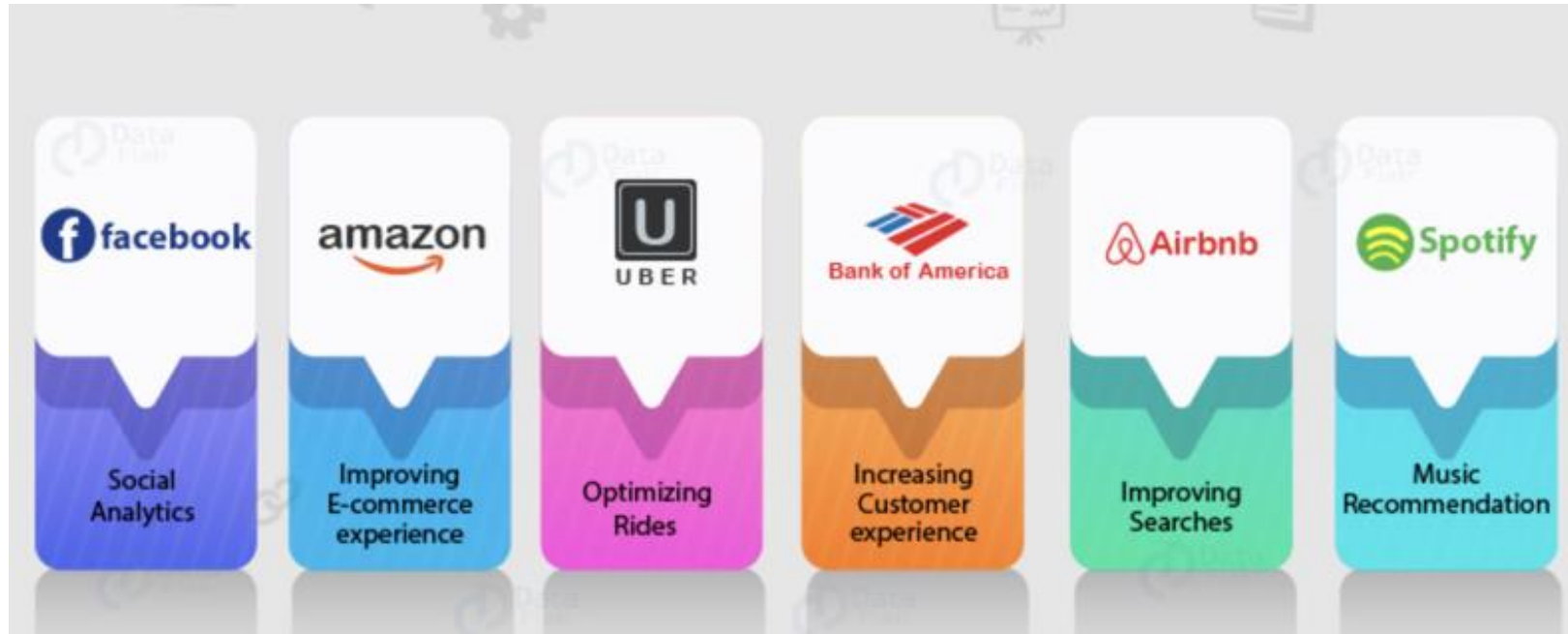


Visualization: Uber movement traffic data



Travel times by days and hours can be predicted based on past data and current conditions

Other «big players» exploiting BI applications






The latest evolution of BI: GPT and Open-AI

- What is GPT: Generative Pre-trained Transformers (a very complex deep learning architecture, will shortly see later in this course) that can be exploited for many natural-language and image-processing tasks such as:
 - Question answering (e.g. [CHAT GPT](#))
 - Content creation (writing documents and reports)
 - Image creations (e.g. [DALL-E](#))
 - Automated generation of code, excel files, etc.

Benefits of GTP/OpenAI systems for business

- **Faster data processing and data analytics:** ChatGPT has the potential to process significant amounts of data in real-time, giving businesses accurate and up-to-date insights that can be used for faster decision-making.
- **Automated creation of natural language reports and summaries:** BI and analytics vendors can also use ChatGPT to help them create reports and summaries faster, simplifying their content so that non-technical users will understand.
- **Improved customer service (last generation chatbots):** Users can also rely on ChatGPT to provide clear, human-like responses to clients and customers, improving their customer service strategy and bolstering the company's overall reputation.




Example 1: summary generation

 Please summarize the text "The shortcuts and built-in snipping tools are only for simple and quick screen capture on HP computers. To take the screenshot experience to the next level, you can download a third-party screenshot tool, such as Awesome Screenshot & Recorder."  

Awesome Screenshot & Recorder is an all-in-one screenshot tool that allows you to take screenshots, record videos, edit and host images/videos, and share files instantly via links. It offers browser extensions for Chrome, Firefox, Edge, and Safari. So, if your HP computer has installed any of the 4 browsers, you can download and use this powerful screenshot tool.

As a handy screen capture tool, it lets you capture screenshots from scrolling web pages, the entire screen or any selected area. Once the screenshot is captured, you can use its built-in annotation tools to resize it, add text/highlights/blur to it, etc. When the image is ready, you can choose to download it locally, upload to the free cloud storage, or quickly share with others."

Please write in English language.

 The article discusses screen capture options available on HP computers. While the built-in shortcuts and snipping tools are useful for quick and simple captures, the article suggests that users who want a more advanced experience should consider downloading third-party tools such as Awesome Screenshot & Recorder.  

Awesome Screenshot & Recorder is an all-in-one tool that allows users to capture screenshots and videos, edit and host images and videos, and instantly share files via

 Regenerate response

Output in

English

Tone

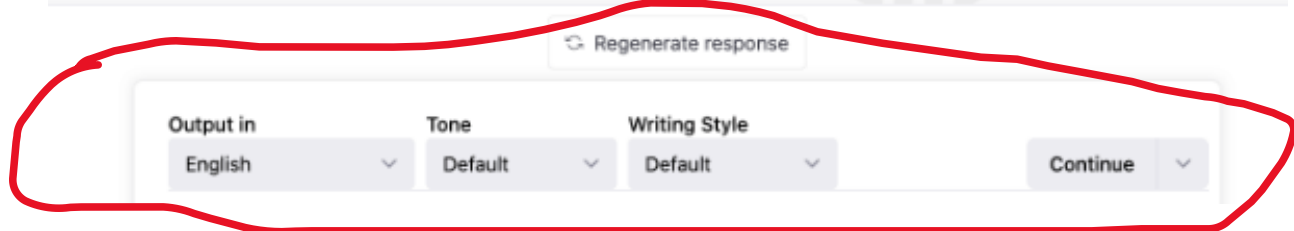
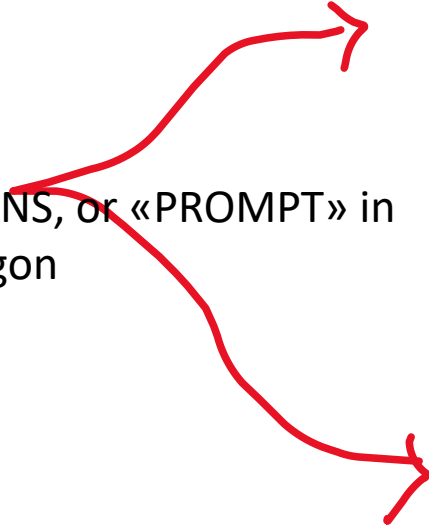
Default

Writing Style

Default

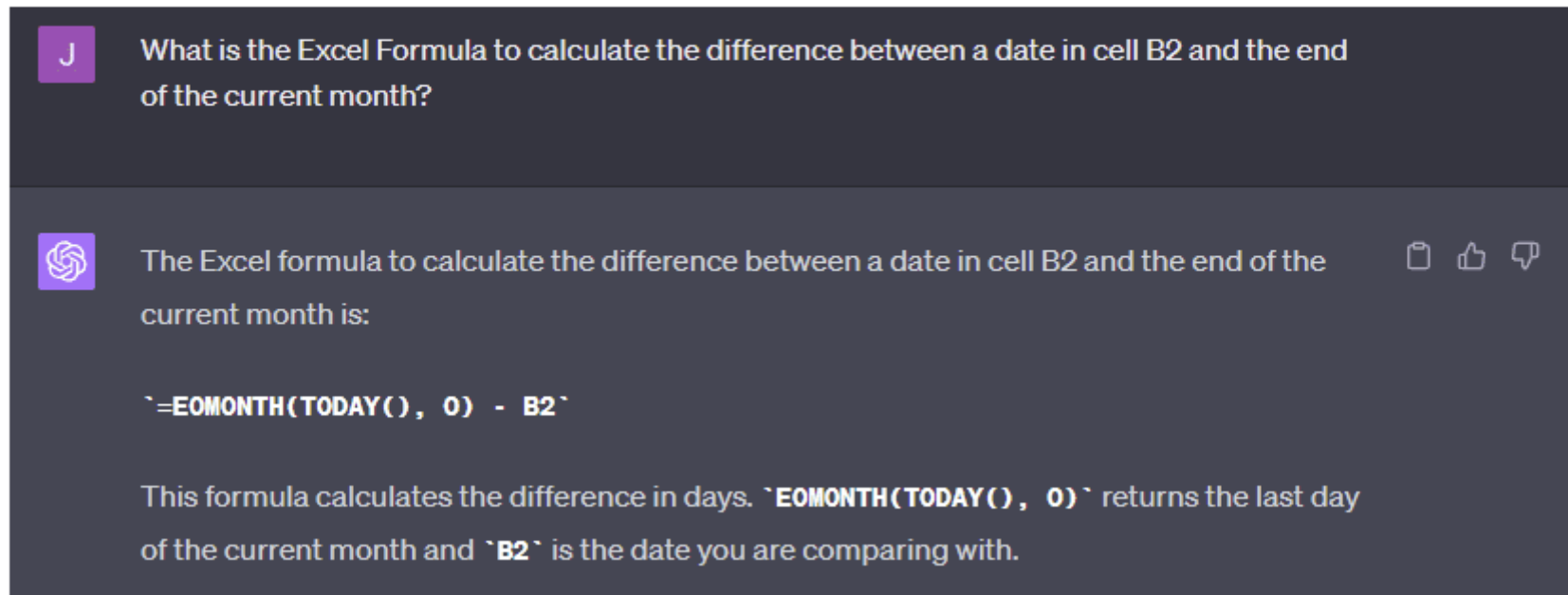
Continue

INSTRUCTIONS, or «PROMPT» in
The GPT jargon



Example 2: generating excel files or code automatically

Example #1. Asking ChatGPT to generate a formula to calculate the days between a date and the end of the current month.



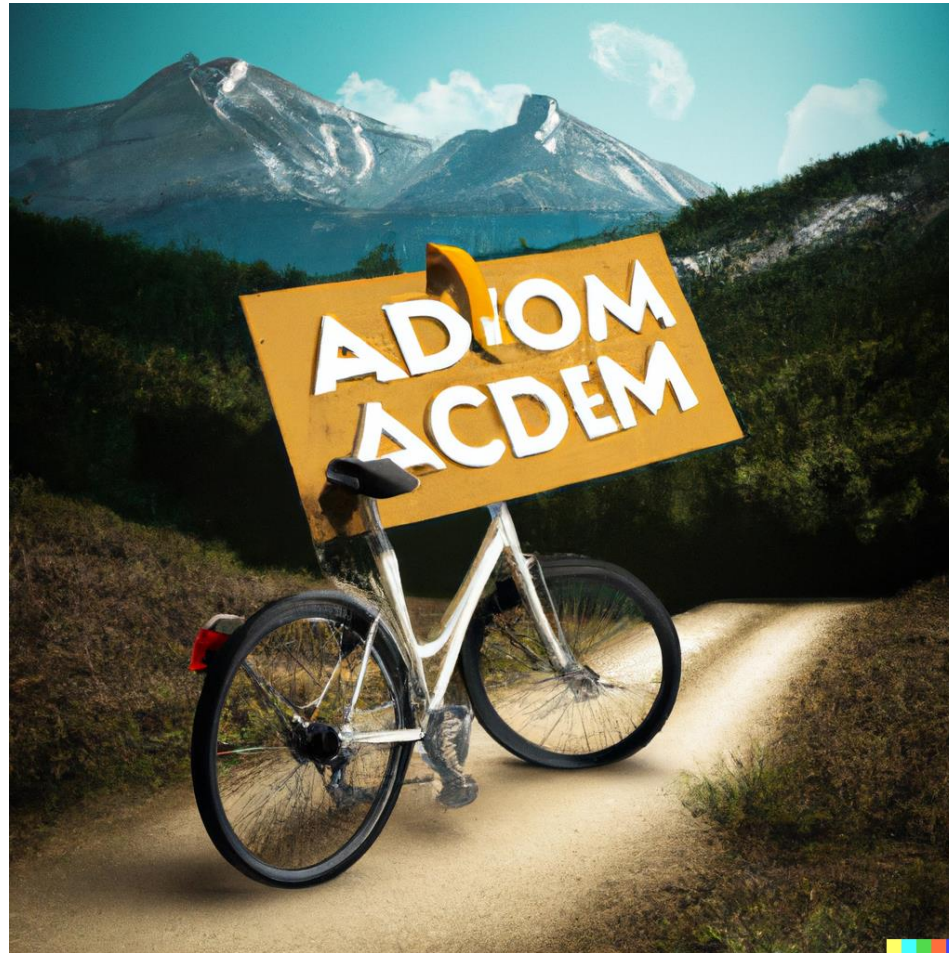
J What is the Excel Formula to calculate the difference between a date in cell B2 and the end of the current month?

The Excel formula to calculate the difference between a date in cell B2 and the end of the current month is:

```
=EOMONTH(TODAY(), 0) - B2
```

This formula calculates the difference in days. `EOMONTH(TODAY(), 0)` returns the last day of the current month and `B2` is the date you are comparing with.

Example 3: automated creation of advertising images



Instruction: advertising image of a mountain bike, digital art