Database Technologies

MS218
Semester 1
Session 2021-22

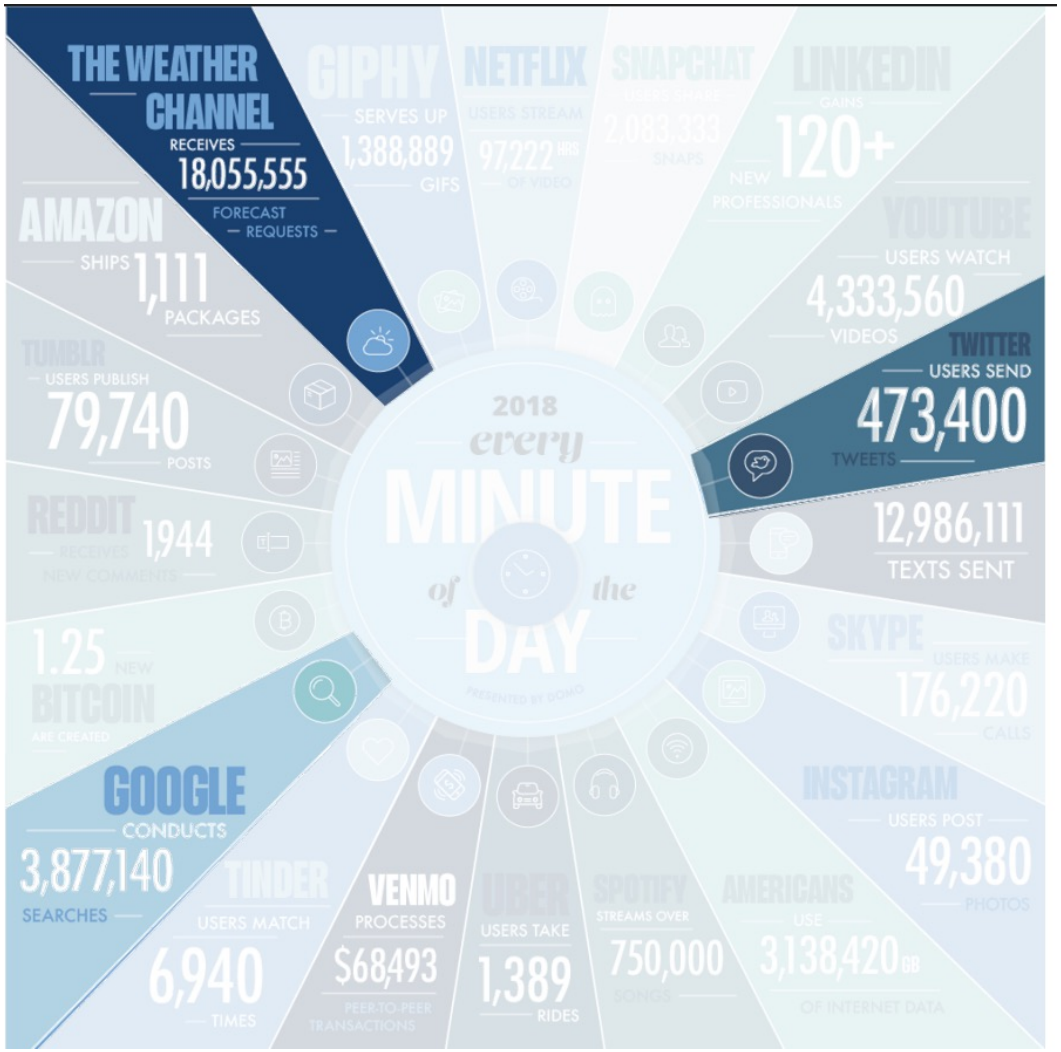
Who: Course Staff

- Your Instructor: Jamal Abdul Nasir
  - Lecturer at BIS, School of Business & Economics
  - Teaching this course 6th time at Undergraduate level
  - Research Area: Data Mining, Distributed Computing, and Machine Learning
  - Office Hours: TBD
  - Office#375
  - Email: jamal.nasir@nuigalway.ie

The world is increasingly driven by data…

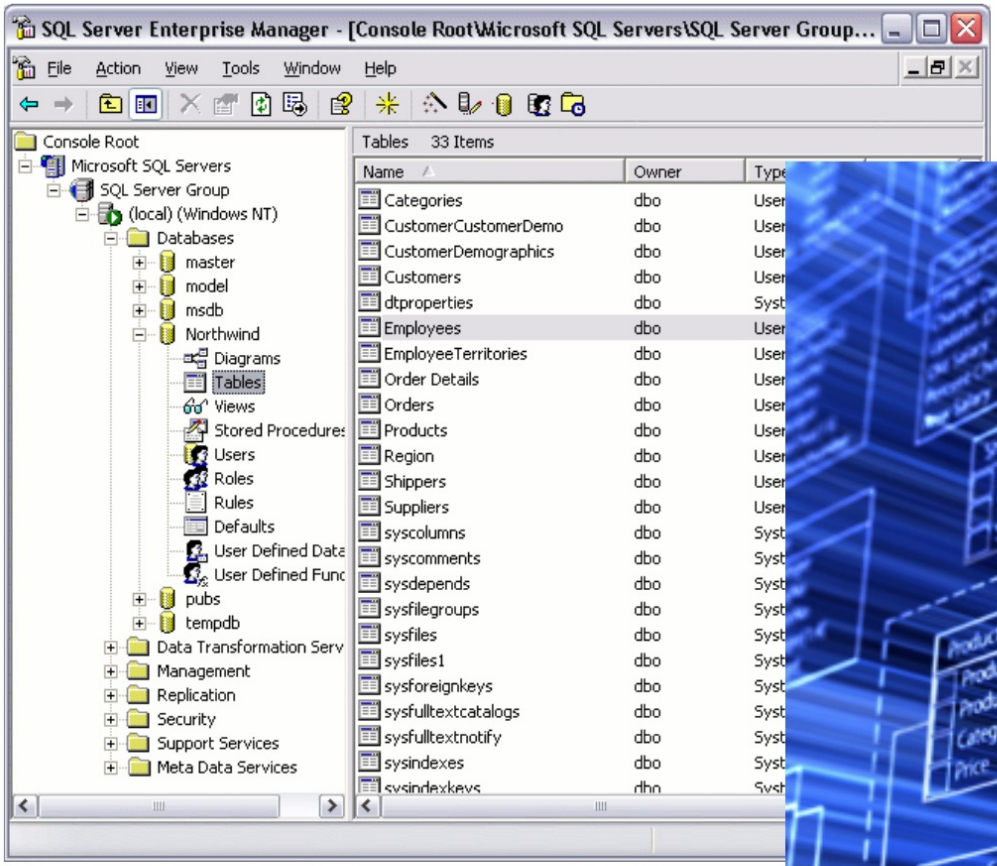
This class teaches **the basics** of how to use & manage data.

Data



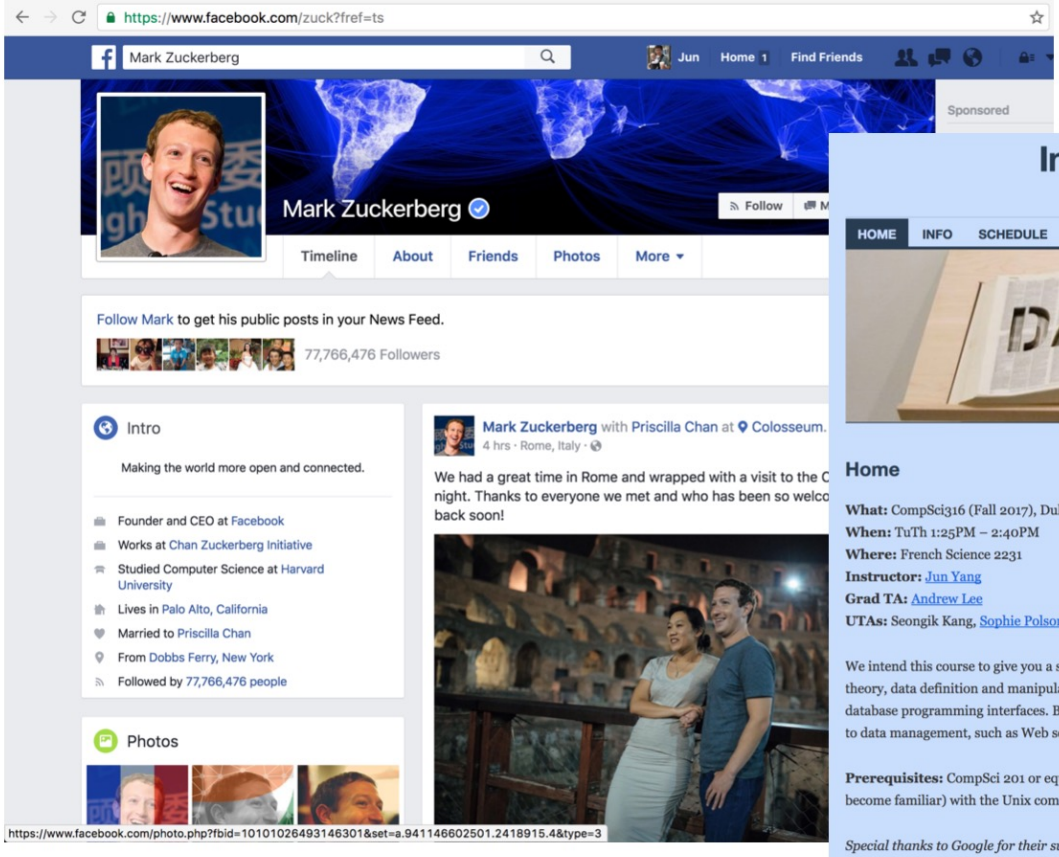
https://www.domo.com/learn/infographic/data-never-sleeps-5

What comes to your mind…



http://www.quackit.com/pix/database/tutorial/dbms_sql_server.gif
http://webstoresltd.com/wp-content/uploads/2013/06/database-design.jpg

But these uses databases too…



Facebook uses MySQL to store posts, for example

WordPress uses MySQL to manage components of a website (pages, links, menus, etc.)

Data → Power

-

… a then little-known British company based in London sent out a press release: "We are thrilled that our revolutionary approach to data-driven communication has played such an integral part in President-elect Trump's extraordinary win," Alexander James Ashburner Nix was quoted as saying. Nix is British, 41 years old, and CEO of Cambridge Analytica. … His company wasn't just integral to Trump's online campaign, but to the UK's Brexit campaign as well.

https://motherboard.vice.com/en_us/article/mg9vvn/how-our-likes-helped-trump-win

Communication

- Website: Blackboard
- Calendar, schedule, policies, labs, links, assignments, etc.
  - Grade book and assignment submissions via BB
- Discussion:
  - Ask and answer questions
  - *ALL* questions on course material should go here
  - Can email to make individual appointments

Course Components and Grading

- **Individual Homeworks** (20% total)
- **Group Project** (20%) Maximum 2 students in a group
- **Final Exam** (60%)
  - Comprehensive

Submission Policy

- Late penalties: 1 day = 10%; 2 day late = 20%; 3 day late = 30%
- You may only submit ONCE with 50% deduction if you are more than three days late.

How to Get an A

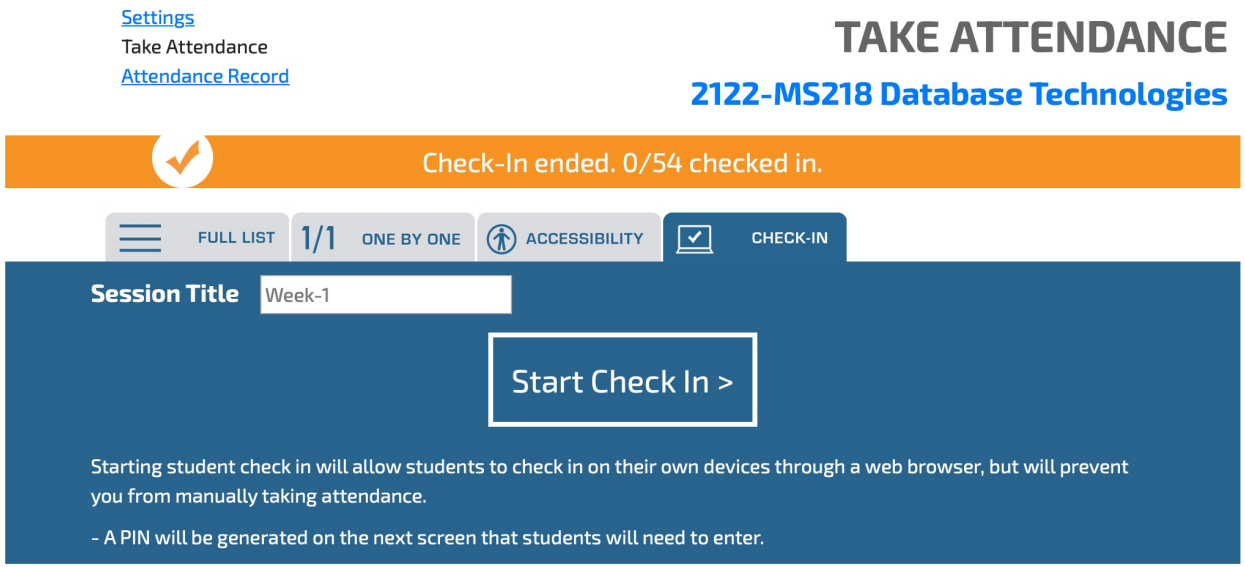
- Attend every class
- Complete your assignments on time
- Reach out to us if you ever feel stuck or overwhelmed
- **Persistence is important**: a lot of things will seem new and confusing at first, but you can figure them out – stick with it and don't give up!
  - You learn best from your mistakes

How to Get an F

- Attend class after few weeks
- Copy your assignments from others
- Never try to understand concepts

Attendance

- I dislike mandatory attendance… but in the past we noticed…
  - People who did not attend, failed in the course
  - People who did not attend were less happy with the course



Course Outline

Units	Readings
Week 1: Database Introuction	Chapter-1 (RS)
Week 2: Conceptual Modelling • ERD	Oracle Notes
Week 3: UIDs • UIDs • Normalization	Oracle Notes
Week 4: Conceptual Modelling and Mapping • Arcs, and Recursive Modelling • Mapping	Oracle Notes
Week 5: SQL • Relational Algebra • SQL	Chapter-6 (RS) Chapter-4 (RS)
Week 6: SQL • Functions	Chapter-4 (RS)
Week 7: SQL • Functions	Oracle Notes
Week 8: SQL • Groups	Chapter-5 (RS)
Week 9: SQL	Oracle Notes
Week 10: Transaction Processing • Transaction Processing	Chapter-21 (RS)
Week 11: Adnavced Databases • **Databases in new era**	
Week 12: Course Review	

So what is a database?

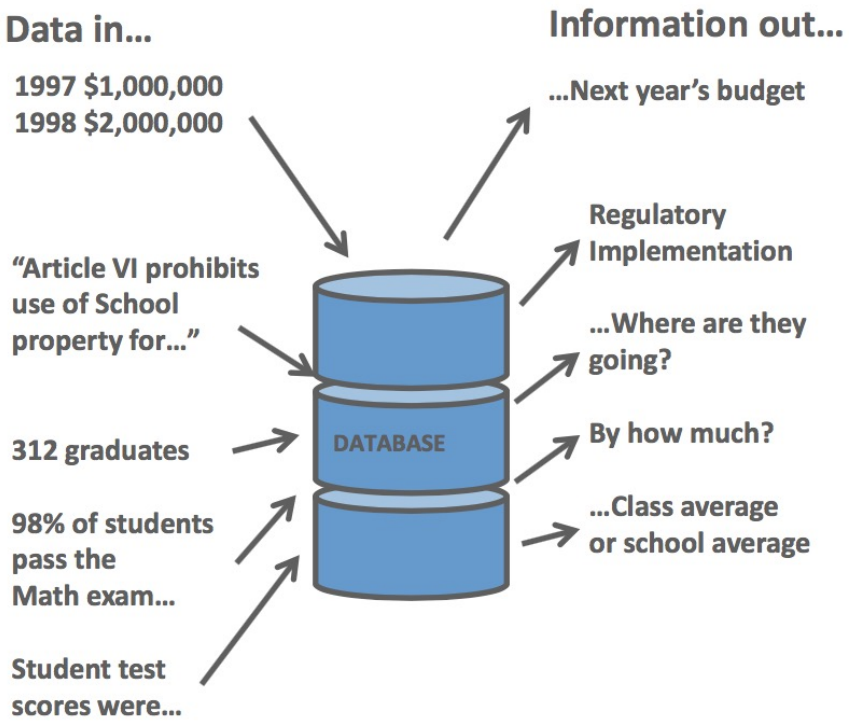
- Database: an organized body of related information (From Oxford Dictionary)
- All kinds of information (school records, mobile telephone records, ring tone downloads, grocery purchases) are stored in databases
- Miniworld or universe of discourse (UoD)

Data vs. Information

- Data: Unprocessed material
- Information: knowledge, intelligence, a particular piece of data with a special meaning or function.
- Information is often the result of combining, comparing, analyzing or performing calculations on data.

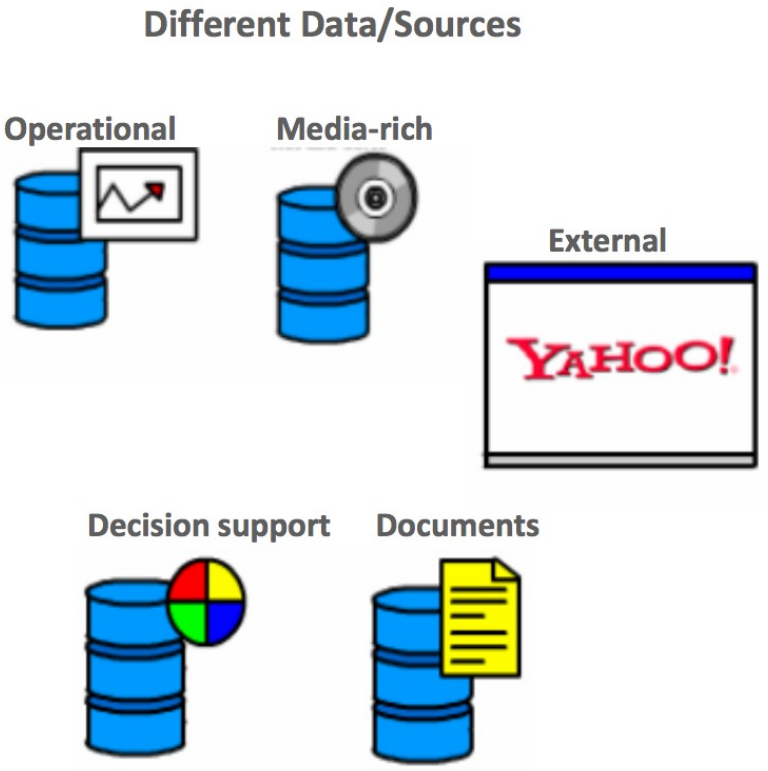
Data vs. Information

- The Oracle database software will transform recorded/stored data and statistics into useful pieces of information.
- Data: Each student's test score is one piece of data.
- Information: The class's average score or the school's average score.



Databases

- Within most modern databases, you can store and retrieve a wide variety of data and documents.
- Inside the database, data is stored in its “raw” form.
- When this raw data is queried or retrieved, it is transformed into more useful information.



So what is a database system?

- Database system, DataBase Management System (DBMS): a software system that facilitates the creation and maintenance and use of an electronic database

Examples of DBMS

DBMS Rankings

363 systems in ranking, December 2020

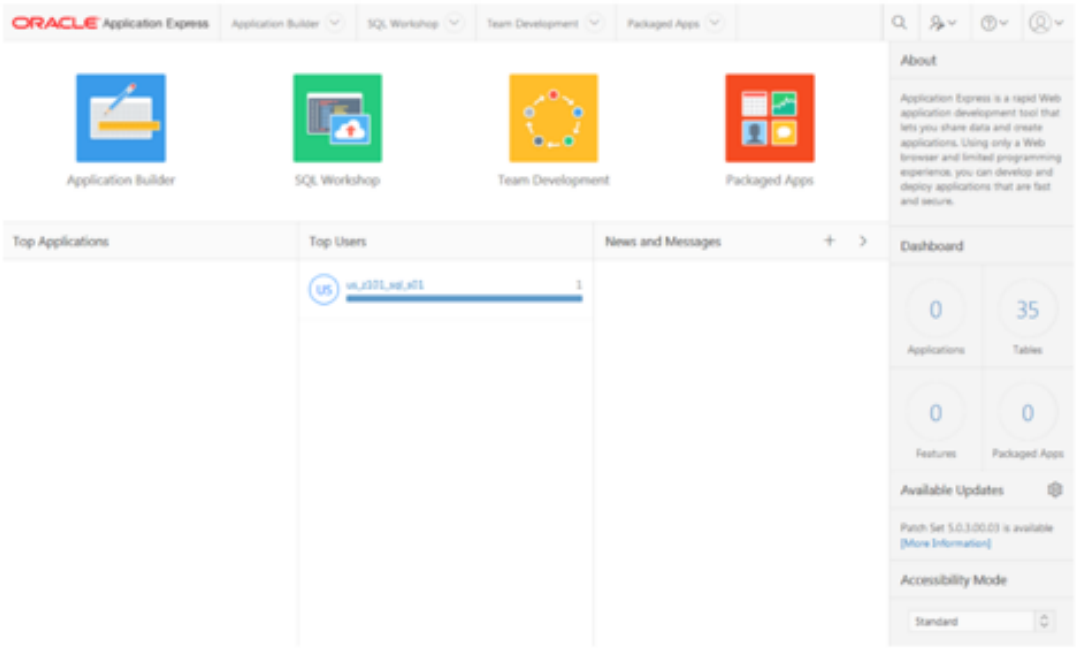
Rank Dec 2020	Rank Nov 2020	Rank Dec 2019	DBMS	Database Model	Score Dec 2020	Score Nov 2020	Score Dec 2019
1.	1.	1.	Oracle	Relational, Multi-model	1325.60	-19.40	-20.80
2.	2.	2.	MySQL	Relational, Multi-model	1255.45	+13.81	-20.21
3.	3.	3.	Microsoft SQL Server	Relational, Multi-model	1038.09	+0.45	-58.11
4.	4.	4.	PostgreSQL	Relational, Multi-model	547.57	-7.49	+44.20
5.	5.	5.	MongoDB	Document, Multi-model	457.73	+3.90	+36.61
6.	6.	6.	IBM Db2	Relational, Multi-model	160.43	-1.19	-10.91
7.	7.	↑ 8.	Redis	Key-value, Multi-model	153.63	-1.79	+7.39
8.	8.	↓ 7.	Elasticsearch	Search engine, Multi-model	152.49	+0.94	+2.24
9.	9.	↑ 11.	SQLite	Relational	121.68	-1.63	+1.32
10.	10.	10.	Cassandra	Wide column	118.84	+0.09	-1.87
11.	11.	↓ 9.	Microsoft Access	Relational	116.74	-0.50	-12.73
12.	12.	↑ 13.	MariaDB	Relational, Multi-model	93.61	+1.31	+6.82
13.	13.	↓ 12.	Splunk	Search engine	87.00	-2.71	-3.53
14.	14.	↑ 15.	Teradata	Relational, Multi-model	73.83	-1.77	-4.66
15.	15.	↓ 14.	Hive	Relational	70.27	+0.01	-15.78
16.	↑ 17.	↑ 25.	Microsoft Azure SQL Database	Relational, Multi-model	69.49	+2.50	+41.60
17.	↓ 16.	↓ 16.	Amazon DynamoDB	Multi-model	69.12	+0.23	+7.49
18.	18.	18.	SAP Adaptive Server	Relational	54.88	-0.51	-0.66
19.	↑ 20.	↑ 22.	Neo4j	Graph	54.64	+1.10	+4.08
20.	↓ 19.	20.	SAP HANA	Relational, Multi-model	52.50	-1.08	-1.67

https://www.endava.com/en/blog/Engineering/2021/Following-the-patterns-the-rise-of-neo4j-and-graph-databases

Which tool will we use?

Oracle

- Access to their cloud based server (Oracle Application Express, APEX) and softwares.



APEX

- Oracle Application Express (APEX) is a low-code development platform that enables you to build stunning, scalable, secure apps, with world-class features, that can be deployed anywhere.

What do you want from a DBMS?

- Keep data around (**Persistent**)
- Answer questions (**Queries**) about data
- **Update** data

Example: A traditional banking application

- **Data**: Each account belongs to a branch, has a number, an owner, a balance, ...; each branch has a location, a manager, ...
- **Persistency**: Balance can't disappear after a power outage
- **Query**: What's the balance in John's account? What's the difference in average balance between CS and Physics department accounts?
- **Update**(Modification): John withdraws 1000€; charge accounts with lower than 10000€ balance, a 10€ fee