



# 00 - Introduction

## Data Science and Management

Corso di Laurea Magistrale in Ingegneria  
Gestionale

Marco Mamei, Natalia Hadjidimitriou

{marco.mamei, selini}@unimore.it

- Course Program
- Course Organization
- Exam
- Software
- Books
- Calendar

# Course Programme

---

## Data Science

- Some statistics background
- Data exploration and visualization
- Machine learning and data mining
- ML and Optimization

## Data Management

- Data warehouses
- Data lake
- Big data
- ERP

## YouTube Channels for last years

- [https://www.youtube.com/playlist?list=PLm63\\_RCCYVFt2PymuODsGG555FkFM2dLw](https://www.youtube.com/playlist?list=PLm63_RCCYVFt2PymuODsGG555FkFM2dLw)

# Course Organization



Class lectures and code sessions

- **All the teaching material available on Moodle**

Schedule

- Tuesday, 14:00-18:00
- Thursday, 14:00-17:00 / 18:00

# Course Organization



## Additional notes

- Total amount of hours: **81**
- The course should finish **by mid December**

# Course Organization



## Office hours

- Online, using Teams or Google Meet
- Please send me an e-mail in advance
- You can also ask me (quick) questions via e-mail

# Course Organization



Please **subscribe** to the Moodle website!

- I will send communications through the platform
- Changes in lecturing timeline, new slides, etc...

# Exam



The exam will consist of **two parts**:

- Project (weight ~50%)
- Oral exam (weight ~50%)



# Exam



About the project (more info throughout the course):

- You can work in teams of **up to 3 people**
- We will assign a project to each student / team, or you can propose a topic
- The project requires:
  - **Code:** implementation or experimentation
  - **Report:** 4-pages document in english presenting the problem, the solutions and the results. It is important to **motivate** the choices you made
  - **Presentation:** slides for 10 mins project presentation

Oral exam will start with project discussion followed by questions about course's topics

Dates

- Dates to be confirmed:
- January, February, April, June, September, November

Posso usare AI Generativa ed LLM per il corso ? **SI**

Come potremmo usarla in modo strategico ?

## Apprendimento e comprensione

- Ottenere spiegazioni di frammenti di codice  
*Esempio: “Spiegami riga per riga cosa fa questa funzione in Python.”*
- Migliorare e ottimizzare lo stile del mio codice  
*Esempio: “Rendi il mio codice più leggibile e ben documentato.”*
- Chiarire concetti complessi (oltre al codice)  
*Esempio: “Spiegami la differenza tra RandomForest e GradientBoosting.”*
- Comprendere meglio slide e contenuti tecnici  
*Esempio: “Riassumi per me i punti chiave di questa lezione sulle Reti Neurali*

## Sviluppare competenze

- Rafforzare il pensiero critico  
*Esempio: “Quali sono i limiti di questa soluzione al problema?”*
- Potenziare le capacità di problem solving  
*Esempio: “Mostrami tre modi diversi per implementare un algoritmo di ordinamento.”*
- Allenarmi a trovare soluzioni alternative tramite brainstorming  
*Esempio: “Suggeriscimi strategie per ottimizzare le prestazioni di questo programma.”*

## Preparazione e pratica

- Simulare esercitazioni pratiche  
*Esempio: “Genera un esempio simile a questo, ma leggermente più difficile.”*
- Prepararmi a esami e presentazioni con riassunti e quiz generati dall'IA  
*Esempio: “Crea 5 domande a risposta multipla sugli argomenti di questa lezione.”*

**L'AI è uno strumento di supporto e non deve fare il lavoro al posto vostro, altrimenti rischiate di non apprendere davvero.**

# Quali sono i “rischi” ?

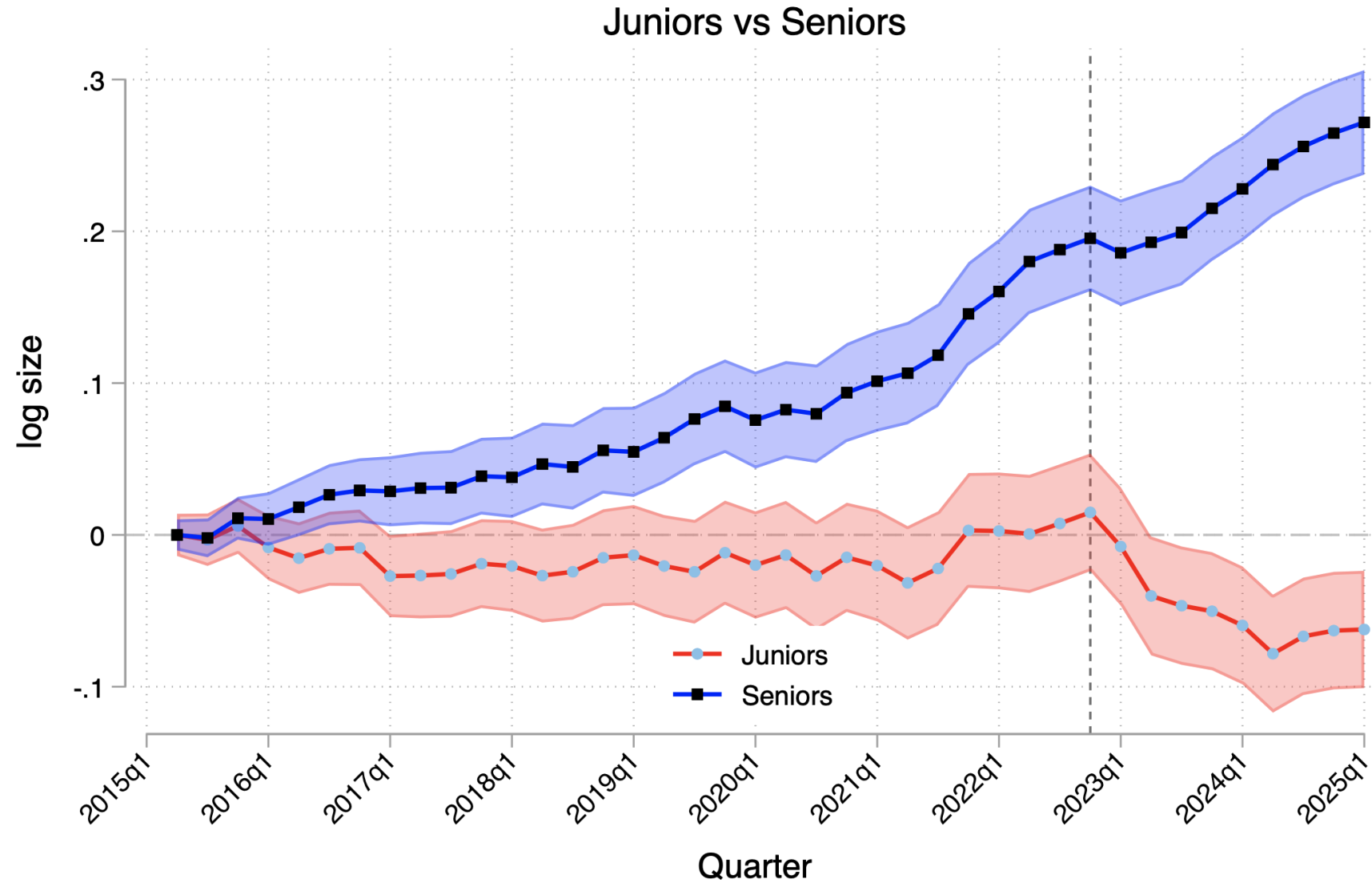
---

- **Perdita di apprendimento reale:** Se l'IA fa tutto il lavoro, lo studente non sviluppa competenze pratiche né capacità di ragionamento.
- **Dipendenza dall'IA:** Rischio di diventare incapaci di risolvere problemi senza l'aiuto dello strumento.
- **Superficialità nella comprensione:** Le risposte possono sembrare corrette ma non sempre sono approfondite: si rischia di “sapere a memoria” senza capire davvero.
- **Errori e allucinazioni:** L'AI può inventare informazioni, produrre codice non funzionante o concetti inesatti.
- **Mancanza di pensiero critico:** Accettare passivamente la risposta senza verificarla impedisce di sviluppare senso critico.
- **Ridotta capacità di problem solving:** Non allenandosi su problemi complessi, si rischia di non saper affrontare sfide nuove o non standard.

# Quali sono i “rischi” ?

- Il processo di apprendimento è intrinsecamente legato a **provare, sbagliare e correggere**. Solo così si sviluppano competenze profonde, pensiero critico e problem solving, elementi indispensabili non solo per eccellere nel percorso di studi, ma anche per diventare professionisti capaci e insostituibili nel mercato del lavoro.
- **Rischi per un Ing. Junior che delega tutto all'AI**
  - Competenze limitate
  - Facile da sostituire da strumenti automatici (l'AI va veloce)
  - Dipendenza dall'IA anche per compiti semplici
- **L'AI è un moltiplicatore di capacità, non un sostituto. Chi la usa bene diventa più prezioso sul lavoro; chi la usa male rischia di ridurre la propria rilevanza.**
- **Puntare su competenze e qualità**

# Il Mercato del Lavoro Cambia !



- La valutazione del codice e dei progetti **non richiede nozioni a memoria**, ma **comprensione e responsabilità**.
- Ogni studente dovrà essere in grado di: Spiegare **ogni porzione di codice** cosa viene fatto per risolvere quale problema e con quale approccio
- Descrivere e motivare **l'uso di librerie e strumenti**
- Motivare le **scelte progettuali e soluzioni implementate**
- Anche se il codice funziona correttamente, **se non sarà in grado di spiegarlo, l'esame non sarà superato.**
- L'obiettivo: sviluppare **autonomia, responsabilità e padronanza del proprio lavoro.**

# Software

---

For the data science part, we will use Python 3+

Available for Windows, Linux, MacOSX

You can use any IDE you like: during the course - I will use VS Code.

<https://www.python.org/downloads/>

<https://code.visualstudio.com/download>

<https://github.com/settings/education/benefits> (copilot)

For the data management part of the course, we will use Postgress

<https://www.postgresql.org/>



# Books



- J. VanderPlas. Python Data Science Handbook, O'Reily
- A. Géron. Hands-On Machine Learning with Scikit-Learn, Keras, and Tensorflow: Concepts, Tools, and Techniques to Build Intelligent Systems, O'Reily
- R. Kimball, M. Ross. The Data Warehouse Toolkit, Third Edition, Wiley
- M. Golfarelli, S. Rizzi. Data Warehouse: Teoria e Pratica della Progettazione, Seconda Edizione. McGraw-Hill