

# Progetto finale (= esame)

Introduzione alla Data Science

Nicoletta Noceti

# Come ci organizziamo

- **Oggi parliamo dei progetti**
  - Calendario
  - Dataset
  - Voto
- **Ricordate che progetto=esame, quindi va fatto singolarmente**

# Calendario

11/04	12/04	13/04	14/04	15/04	16/04	17/04
18/04	19/04	20/04	21/04	22/04	23/04	24/04
25/04	26/04	27/04	28/04	29/04	30/04	01/05
02/05	03/05	04/05	05/05	06/05	07/05	08/05
09/05	10/05	11/05	12/05	13/05	14/05	15/05
16/05	17/05	18/05	19/05	20/05	21/05	22/05
23/05	24/05	25/05	26/05	27/05	28/05	29/05
30/05	31/05					

**Assegnazione progetto**

**Sviluppo del progetto**

**Consegna del progetto**

**Peer review**

**Presentazione del progetto**



**Dataset**

# Wordle Tweets (1)

Wordle tweets dataset to help one analyze the distribution of others attempt



<https://www.kaggle.com/datasets/vora1011/wordletweets>

# Superstore dataset (2)

Dataset containing Sales & Profits of a Superstore



<https://www.kaggle.com/datasets/vivek468/superstore-dataset-final>

# TED Talks (3)

This datasets is having all the ted talks available on Ted.com



<https://www.kaggle.com/datasets/ashishjangra27/ted-talks>

# Amazon Top 50 Bestselling Books 2009 – 2022 (4)

The dataset contains data from bestselling books on Amazon



<https://www.kaggle.com/datasets/chriskachmar/amazon-top-50-bestselling-books-2009-2022>



# Voice Call Quality Customer Experience (5)

Based on crowdsource data collected through MyCall app



<https://www.kaggle.com/datasets/arnavr10880/voice-call-quality-customer-experience-india>

# 9000+ Movies Dataset (6)

Movies Dataset sorted by its popularity for Recommender Systems



<https://www.kaggle.com/datasets/disham993/9000-movies-dataset>

# Google Play Store Apps (7)

Web scraped data of 10k Play Store apps for analysing the Android market.



<https://www.kaggle.com/datasets/lava18/google-play-store-apps>

## Goodreads-books (8)

Comprehensive list of books listed in goodreads



<https://www.kaggle.com/datasets/jealousleopard/goodreadsbooks>

# Video Game Sales (9)

Video Games categorized by Sales



<https://www.kaggle.com/datasets/rishidamarla/video-game-sales>

# Music Dataset : 1950 to 2019 (10)

Lyrics and Metadata from 1950 to 2019



<https://www.kaggle.com/datasets/saurabhshahane/music-dataset-1950-to-2019>



# Top 50 Spotify Songs – 2019 (11)

Top 50 songs listened in 2019 on spotify



<https://www.kaggle.com/datasets/leonardopena/top50spotify2019>

# Analisi del movimento



Azioni registrate:

1. Toccare la bottiglia
2. Toccare il cubo di Rubik
3. Spostare la bottiglia
4. Spostare il cubo di Rubik
5. Spostare il pennarello
6. Bere dal bicchiere
7. Mangiare una patatina
8. Aprire la bottiglia
9. Igienizzarsi le mani
10. Giocare con il cubo di Rubik

- Abbiamo acquisito un piccolo dataset di video in cui un soggetto esegue una lista di azioni
- 16 partecipanti x 10 azioni x 2 video = 320 video, ossia circa 60K frames
- Abbiamo estratto una descrizione della posa del partecipante (skeleton) e della direzione della testa in 3D
- Dalla posa possiamo estrarre misure del movimento di ogni punto dello skeleton → possiamo costruire diversi dataset con caratteristiche diverse





**E' il momento di votare**

# Scegliete due dataset specificando il numero tra parentesi

[https://forms.office.com/Pages/ResponsePage.aspx?id=g2\\_TbAlaLUSXLyZwy16bGnDaeTqlUhZDnnAC7aDi2OZUQjFHQzc4RDlTVEVXVWg4MEg4MlEwNTJYSC4u](https://forms.office.com/Pages/ResponsePage.aspx?id=g2_TbAlaLUSXLyZwy16bGnDaeTqlUhZDnnAC7aDi2OZUQjFHQzc4RDlTVEVXVWg4MEg4MlEwNTJYSC4u)

# UniGe

---

