

Corso di Statistica Sociale

CORSO DI LAUREA: SCIENZE DELL'EDUCAZIONE

DOCENTE: FRANCESCO SANTELLI

Orari e contatti

Contatti docente: Email fsantelli@units.it

[Corso: 305SF - STATISTICA SOCIALE 2021 \(units.it\)](#)

Inizio Lezioni: 12/10/2021

Giorno: martedì e giovedì. Orario: 10.00 – 13.00. Dieci lezioni da 3 ore ciascuna. Totale 30 ore, corso di 6 CFU.

Le lezioni si svolgeranno in **presenza** ma verranno registrate (disponibili su **Teams**, codice iqrn7zs)

Cosa devo sapere prima di iniziare?

- Nessun corso prima di questo è necessario. Potrebbe anche essere il vostro primo esame!

- Le pochissime conoscenze necessarie sono di ambito matematico:

1) Le 4 operazioni elementari

2) I simboli matematici di base

3) Come elevare al quadrato e come fare radice quadrata (con calcolatrice..)

Cosa farò durante il corso?

- Ci saranno lezioni teoriche di stampo più tradizionale
- Seguiranno esercizi con metodo classico, «**carta&penna**»
- Ci saranno fasi di applicazioni pratiche: principalmente utilizzo **excel**.
- Ci saranno dei **mini-compiti** di settimana in settimana (non li correggerò necessariamente, ma mi aspetto che ci proviate e che mi chiediate la volta dopo cosa non avete capito).

Organizzazione tipo di una lezione

Inizio ore 10

Prima parte: Circa un'ora e mezza di lezione teorica, dove cercherò anche di svolgere anche alcuni esercizi alla lavagna e voi mi seguirete (mi auguro almeno) «**carta&penna**»

Pausa di 15-20 minuti

Seconda parte più interattiva: ci saranno esercizi da svolgere da parte vostra, questa volta con il supporto di **excel** (cercate di installarlo sui vostri pc e di avere un portatile con voi, magari organizzatevi a gruppi)

Fine ore 13

Come verrò valutato?

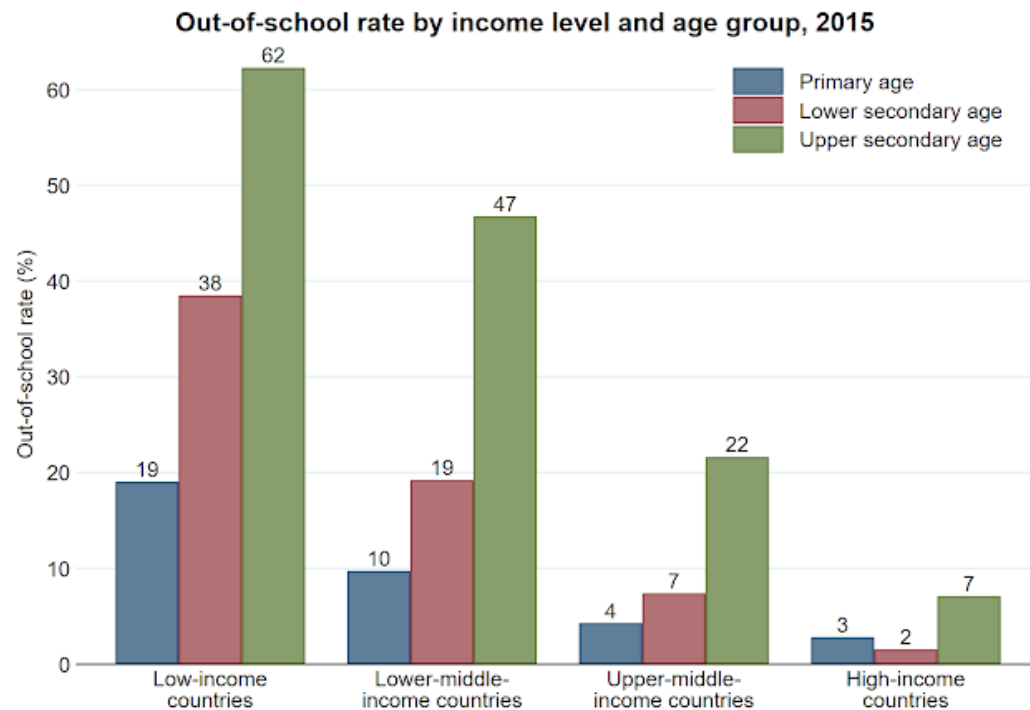
- Obbligatorio: verifica finale scritta (opzione per prova orale). Domande di teoria ed esercizi. Si può utilizzare **calcolatrice** e portare un piccolo **formulario**. Valutazione massima: 30
- Facoltativo: stesura di un mini report a gruppi. E' caldamente consigliato, ma non obbligatorio. Si può incrementare la valutazione dello scritto fino ad un massimo di 4 punti.
- Entrambe le verifiche saranno poco indirizzate al calcolo e maggiormente alla **comprensione**

Perché la statistica?

«**L'educatore socio-pedagogico** è un professionista che opera nell'ambito educativo, formativo e pedagogico, in rapporto a qualsiasi attività svolta in modo formale, non formale e informale, in prospettiva di life long learning, in una ottica di crescita personale e sociale»

- Un educatore socio-pedagogico informato deve **aggiornarsi**, conoscere le nuove tecniche di insegnamento e capire le possibili difficoltà di particolari categorie di persone secondo i più recenti studi
- Molto spesso tali studi vengono effettuati tramite **raccolta dati** e/o **esperimenti**: in entrambi i casi è coinvolta la statistica
- Anche se non sarete voi a condurre sperimentazioni sulla didattica, dovrete comunque essere in grado di **interpretare i risultati** ottenuti da altri!

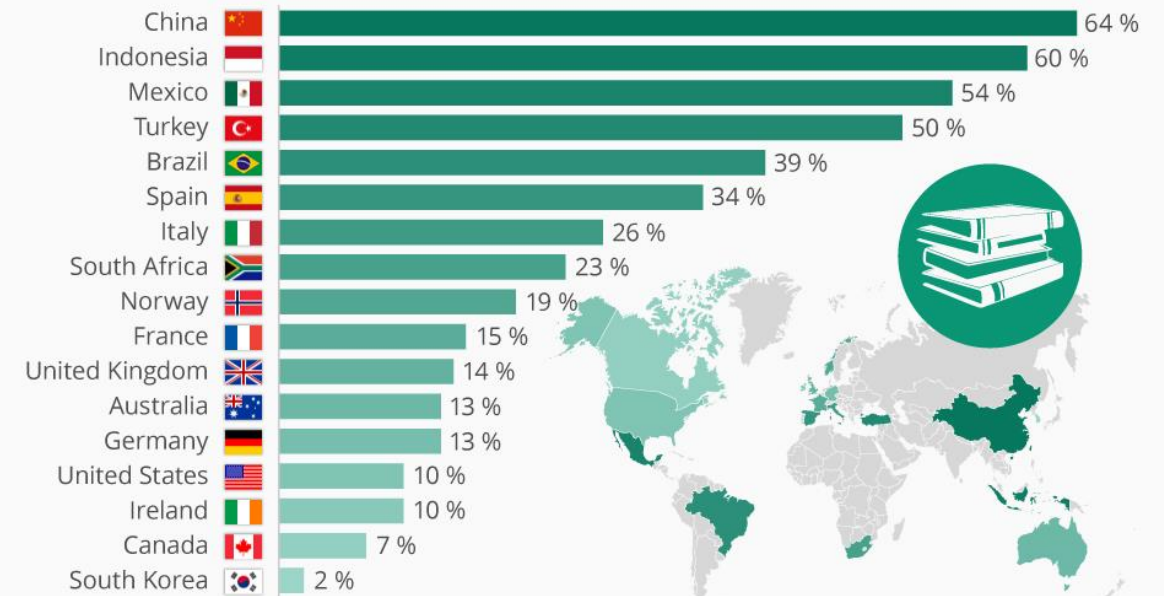
Esempi di utilizzo della statistica in *education (1)*: dati di tipo macro



Friedrich Huebler, huebler.blogspot.com, July 2017

Where Do People Fail To Attain A High School Qualification?

% of 25-34 year-olds with an education level below upper secondary in selected countries*



*2014 or latest year - upper secondary education corresponds to the final stage of secondary education in most OECD countries.

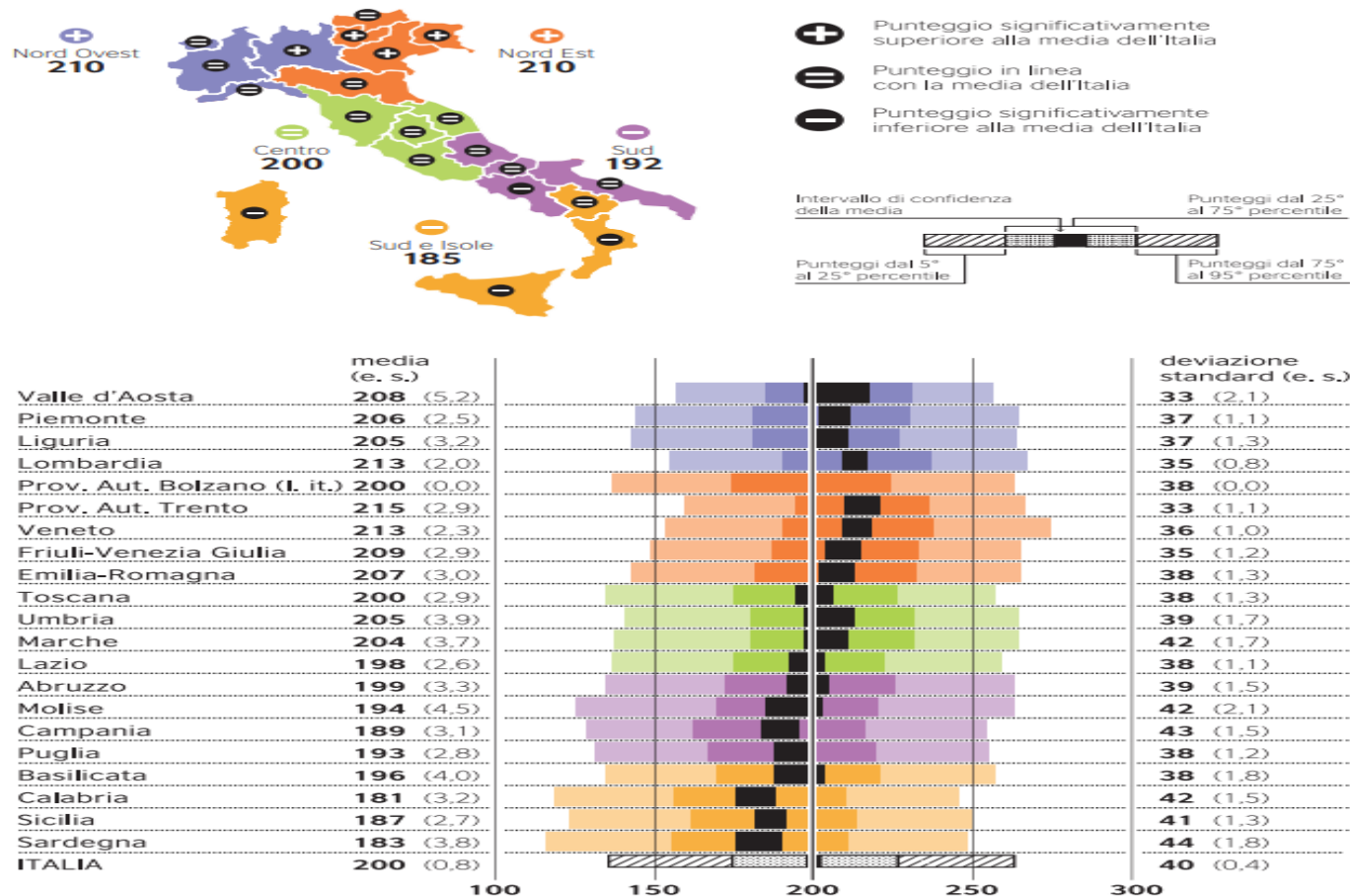


@StatistaCharts

Source: OECD

statista

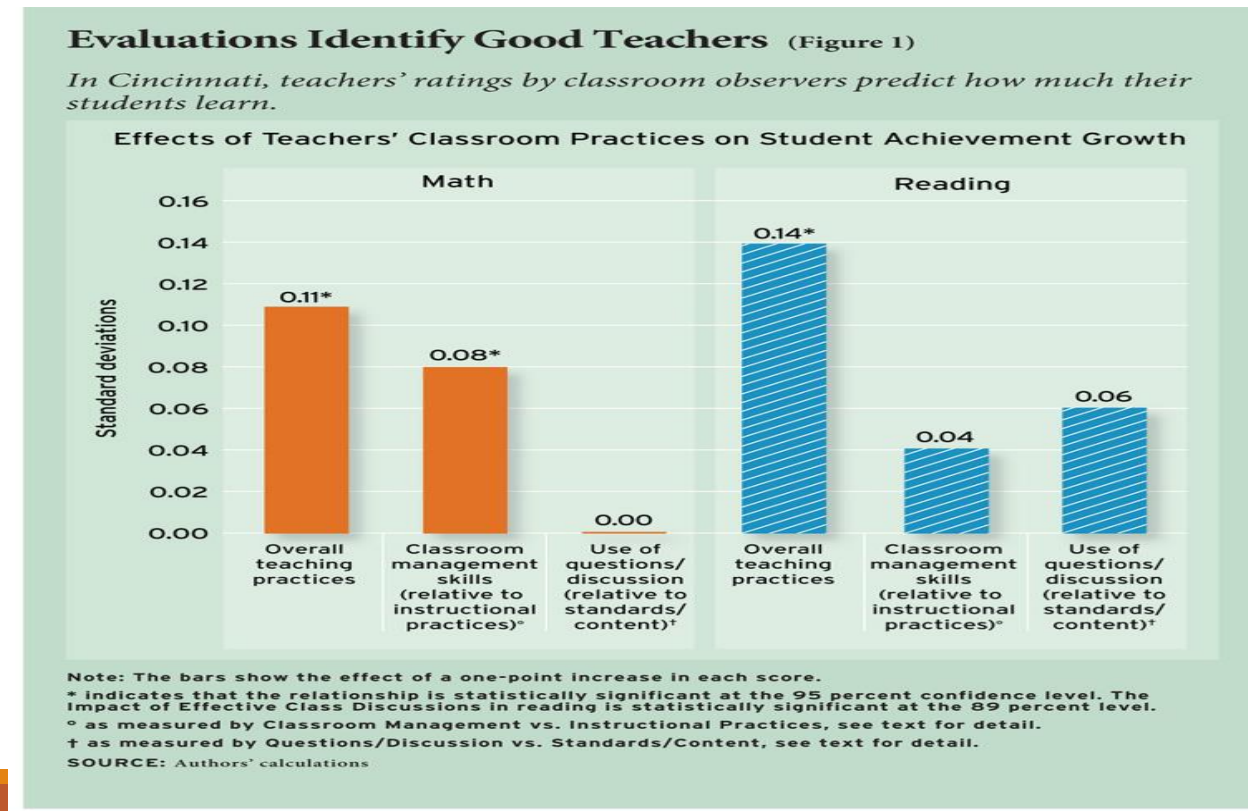
Esempi di utilizzo della statistica in *education (2)*: dati di profitto



Dati **invalsi** sono ampiamente analizzati per valutare eccellenze e criticità nel sistema formativo scolastico.

Esempi di utilizzo della statistica in *education* (3): modelli

Modelli statistici aiutano (ad esempio) a comprendere efficacia di diverse metodologie di insegnamento:



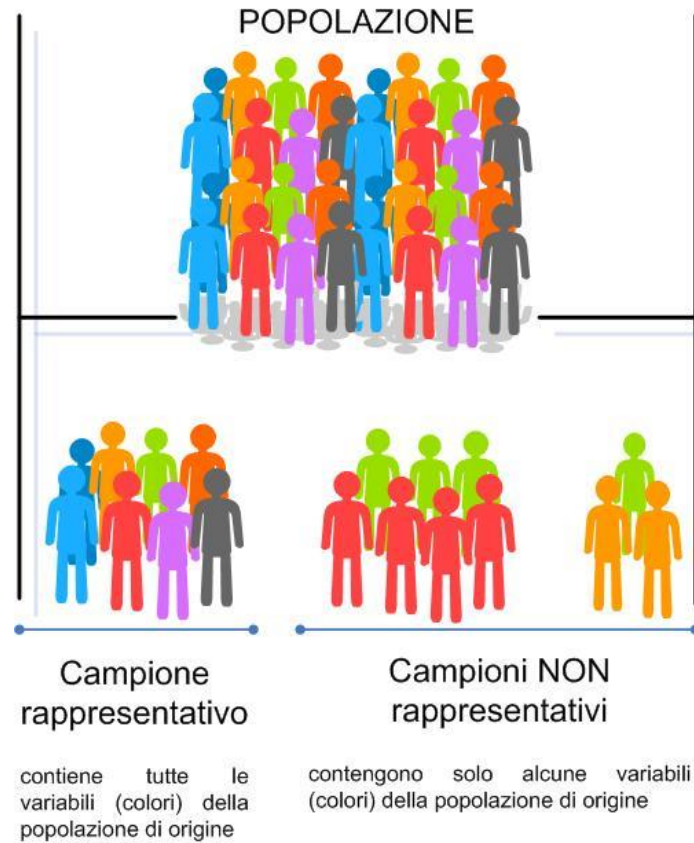
Concetti introduttivi e definizioni principali

Da wikipedia: «La **statistica** è una disciplina che ha come fine lo studio quantitativo e qualitativo di un particolare fenomeno collettivo in condizioni di incertezza o non determinismo, cioè di non completa conoscenza di esso o parte di esso.»

- Fenomeno collettivo: I dati osservati sono riferiti a una moltitudine di persone
- Obiettivo principale: sintetizzare i dati ma allo stesso tempo descriverli
- Ciò può essere fatto con: indici o con grafici o con tabelle
- Il focus del corso sarà su questa parte della statistica, che prende il nome di **statistica descrittiva**

Uso corretto di alcuni termini statistici (1)

- Unità statistica: l'entità su cui è effettuata la rilevazione del dato. Può essere una singola persona, una famiglia, un ateneo, ma anche animali o piante o cose
- Popolazione: l'insieme di tutte le unità statistiche oggetto di analisi
- Se vengono analizzate tutte le unità statistiche (intera popolazione) si parla di **censimento**



- Solitamente, è impossibile analizzare tutta la popolazione! Quanto ci vorrebbe, in termini di costi e tempi, per analizzare tutti i 60 milioni di italiani!?
- Allora dalla popolazione si estrae un **campione**
 - Il campione è un sottoinsieme della popolazione, e la rappresenta in miniatura
 - Se un campione è una fedele riproduzione della intera popolazione allora esso è **rappresentativo**

Rappresentatività: l'esempio della zuppa!

- Poniamo il caso che si voglia studiare una **zuppa** composta da diversi tipi di verdure: patate, carote, fagioli, pomodori ecc.
- Essa rappresenta in termini statistici la **popolazione**
- Per valutare le qualità della zuppa, ne assaggiamo solo un **cucchiaino**.
- Esso è il **campione** della zuppa. Più precisamente, uno dei tanti possibili.
- Se la zuppa è stata «mescolata» bene, allora è probabile che nel cucchiaino che assaggiamo ci sarà un pezzettino di ogni verdura.
- Tale campione (cucchiaino) è quindi **rappresentativo** dell'intera zuppa siccome la rappresenta in miniatura. Se mescolata bene, un cucchiaino è sufficiente per trarre conclusioni sull'intera zuppa
- Se non mescolata bene, prendere anche 3 cucchiaini in superficie potrebbe portarci a conclusioni sbagliate...

RAPPRESENTATIVITA' NON DIPENDE DA NUMEROSITA'!

Uso corretto di alcuni termini statistici (2)

Indagine: ad un gruppo di 100 studenti universitari di Portogruaro viene chiesto di indicare quante volte alla settimana fanno attività sportiva.

- Indagine campionaria: tali studenti non sono tutta la popolazione studentesca ma un sottoinsieme
- Ciò che si rileva è il n° di volte in cui fanno sport a settimana. Tale fenomeno prende il nome di **variabile**. Si indica con **X**
- Le possibili risposte sono: 0, 1, 2...7. Tali possibili risposte (cioè i diversi modi in cui la variabile può presentarsi) prendono il nome di **modalità**. Si indicano con **x_i**
- 20 studenti hanno risposto 0, 15 studenti hanno risposto 1 ecc. Il Numero di risposte associate ad ogni modalità prende il nome di **frequenza (assoluta)**. Si indicano con **F_i**
- Il numero totale di rispondenti (uguale alla somma delle singole frequenze) è detto **numerosità campionaria** e solitamente si indica con **N**

Come strutturare i dati

- Qualunque analisi statistica si regge su una rigorosa **struttura dati** di partenza.

Esempio, voti di una classe ad un compito:

- Metodo «basic», poco utile: dati in serie.

9, 8, 2, 3, 4, 10, 6, 5, 8...7.

- Metodo tabella di frequenza:

I due metodi sono **equivalenti**, nel senso che esprimono esattamente gli stessi dati!

| Voti (xi) | Frequenza (ni) |
|-----------|----------------|
| 2 | 1 |
| 3 | 1 |
| 4 | 3 |
| 5 | 7 |
| 6 | 6 |
| 7 | 6 |
| 8 | 2 |
| 9 | 1 |
| 10 | 1 |



Domanda: Come si passa dai dati in serie alla tabella di frequenza e viceversa?

Esercizio 1

Si considerino i seguenti dati in serie relativi ai voti all'esame di Statistica (9 CFU) di tutti gli studenti del Dispes dell'Università di Trieste residenti in Campania:

18, 19, 25, 25, 30, 28, 19, 20, 22, 23, 24, 25, 28, 24, 23, 18.

Punti da svolgere:

- 1) Che tipo di analisi si sta effettuando? *Campionaria o censuaria?* Discutere
- 2) Si indichi la **numerosità totale**
- 3) Si indichino le **modalità** e le **frequenze assolute** associate
- 4) *Si costruisca la tabella di frequenza*

Tabella di frequenza: oltre le frequenze assolute

| Voti (x_i) | Frequenza (F_i) | Frequenza relativa (f_i) | Frequenza relativa % | Frequenza cumulata (FC_i) | Frequenza relativa cumulata (fc_i) |
|----------------|---------------------|------------------------------|----------------------|-------------------------------|--|
| 18 | 2 | 0,125 | 12,50% | 2 | 0,125 |
| 19 | 2 | 0,125 | 12,50% | 4 | 0,25 |
| 20 | 1 | 0,0625 | 6,25% | 5 | 0,3125 |
| 22 | 1 | 0,0625 | 6,25% | 6 | 0,375 |
| 23 | 2 | 0,125 | 12,50% | 8 | 0,5 |
| 24 | 2 | 0,125 | 12,50% | 10 | 0,625 |
| 25 | 3 | 0,1875 | 18,75% | 13 | 0,8125 |
| 28 | 2 | 0,125 | 12,50% | 15 | 0,9375 |
| 30 | 1 | 0,0625 | 6,25% | 16 | 1 |
| ? | 16 | 1 | 100% | ? | ? |

Alcuni totali di colonna (in rosso) sono evidenziati, altri sono sostituiti da un punto interrogativo. Perché?

Alcuni simboli sono FONDAMENTALI

Sommatoria!!

Modo compatto di scrivere una somma di più elementi. Esempio:

$$\sum_{i=1}^n F_i$$

Questo dove si finisce con la somma, solitamente si arriva fino a n e cioè fino all'**ultimo elemento**

Questo è *cosa* si somma, cioè (in questo caso) i valori delle frequenze assolute

Questo è da dove si inizia a sommare, solitamente da 1 e cioè dal **primo elemento**

Ad esempio, questa sommatoria di tutte le frequenze assolute...
cosa restituisce?

Esercizio 2 (iniziare in aula, presumibilmente finire a casa)

| Voti (x_i) | Frequenza (F_i) | Frequenza relativa (f_i) | Frequenza relativa % | Frequenza cumulata (FC_i) | Frequenza relativa cumulata (fc_i) |
|----------------|---------------------|------------------------------|----------------------|-------------------------------|--|
| 18 | 2 | 0,125 | 12,50% | 2 | 0,125 |
| 19 | 2 | 0,125 | 12,50% | 4 | 0,25 |
| 20 | 1 | 0,0625 | 6,25% | 5 | 0,3125 |
| 22 | 1 | 0,0625 | 6,25% | 6 | 0,375 |
| 23 | 2 | 0,125 | 12,50% | 8 | 0,5 |
| 24 | 2 | 0,125 | 12,50% | 10 | 0,625 |
| 25 | 3 | 0,1875 | 18,75% | 13 | 0,8125 |
| 28 | 2 | 0,125 | 12,50% | 15 | 0,9375 |
| 30 | 1 | 0,0625 | 6,25% | 16 | 1 |
| ? | 16 | 1 | 100% | ? | ? |

- 1) Cercare di scrivere con formula (utilizzando la sommatoria) i risultati ottenuti nelle varie colonne
- 2) Se vi chiedessi la percentuale di studenti che hanno preso un voto al di sotto del 25, quale colonna guardereste?

- 3) Quale è stato il voto con la frequenza più alta? E i voti non presenti in colonna che frequenza hanno?
- 4) Quale è la differenza tra frequenze relative e frequenze relative percentuali?

Per la prossima puntata...

1) Terminare esercizi 1 e 2

2) Installare sul proprio PC Excel. Come detto, potete anche farlo a gruppetti. Ma cercate di avere un pc almeno *ogni 4-5 persone!*

3) Tentare di ripercorrere su **Excel** gli esercizi 1 e 2

4) La prossima lezione inizierà con le **vostre domande** su:

- A) Cosa non avete ben compreso degli argomenti trattati
- B) Quali esercizi/calcoli vi sono risultati difficili