

# Programming Lab

## Lezione 8

*Lavorare veramente 1: creiamo un modello*

Stefano Alberto Russo

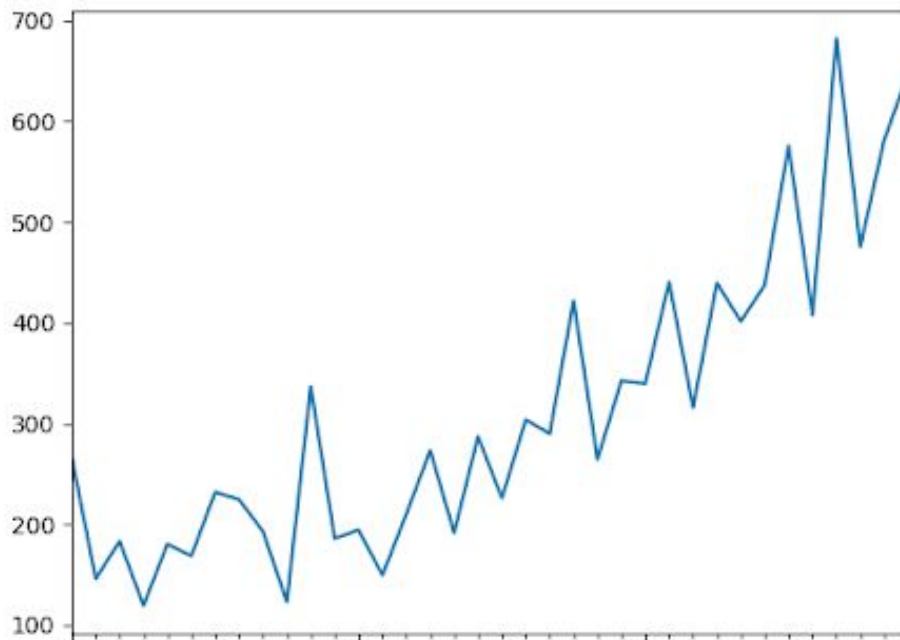
# Cos'è un modello

- “Con modello fisico si intende una rappresentazione concettuale (spesso una semplificazione che ammette una formalizzazione matematica) del mondo reale o di una sua parte, capace di spiegare un determinato fenomeno.”
- “Un modello matematico è una rappresentazione quantitativa di un fenomeno naturale.”
- “Un modello statistico è un modello matematico che incarna un insieme di ipotesi statistiche riguardanti la generazione di dati campione.”

→ *E' una generica ipotesi di come credete che si comporti il mondo reale*

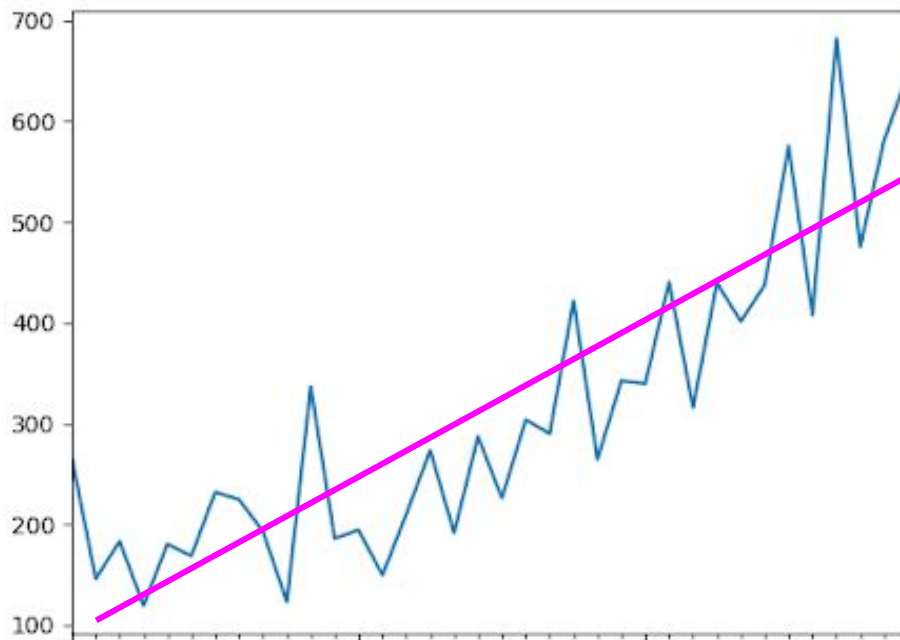
# le vendite degli shampoo

shampoo\_sales.csv



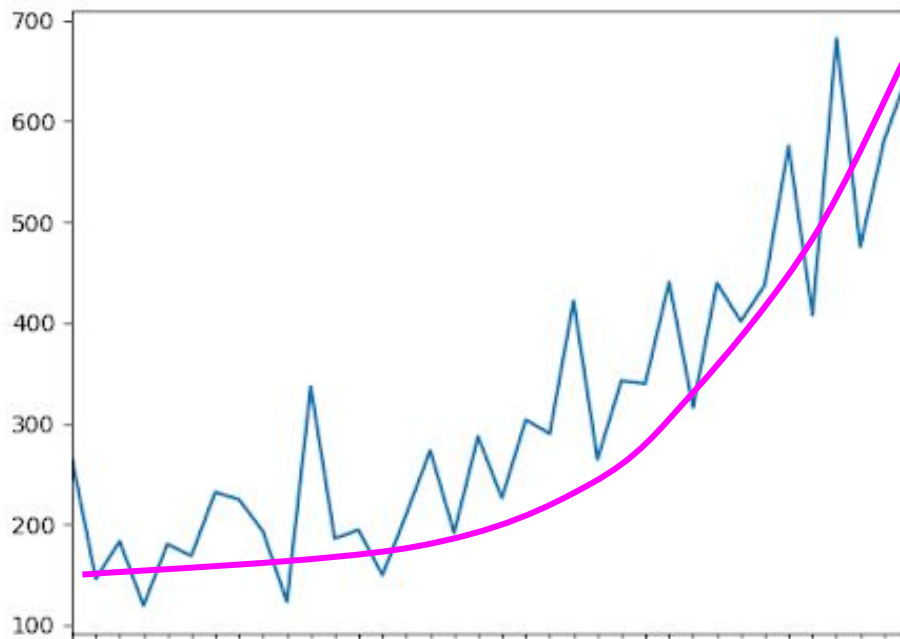
# le vendite degli shampoo

shampoo\_sales.csv



# le vendite degli shampoo

shampoo\_sales.csv



# Perchè mi servono i dati per un modello?

Posso creare anche un modello senza dati.

Per esempio so che le vendite di regali hanno un'impennata sotto le feste. Di quanto aumentano non lo so, ma decido a tavolino del 50%. Questa è la mia ipotesi del modello.

..ma ma se ho dei dati di vendita, posso stimare quanto aumentino le vendite rispetto al mese precedente direttamente dai dati, e magari trovare che aumentano non del 50% ma del 67%. O del 78%.

In tal caso si dice che si fa il ***fit*** del modello (sui dati).

# Cosa me ne faccio di un modello?

Di base, si usano per fare ***predizioni***.

→ Come saranno le vendite dello shampoo il prossimo mese? E il prossimo anno?

→ Quale sarà la parola che viene dopo, se scrivo “ciao come...”?

→ Come si propagherà un’epidemia?

# Il nostro modello per iniziare

Creeremo un modello super-semplificato, tale per cui:

Le vendite dello shampoo al tempo  $t+1$  sono date da

- l'incremento medio negli  $n$  mesi precedenti,
- applicato sulle vendite al tempo  $t$ .



# Esempio

Scelgo di usare 3 mesi per la predizione (**n=3**) e voglio prevedere le vendite a Dicembre (il mio **t+1**).

So che le vendite a Settembre (t-2), Ottobre (t-1) e Novembre (t) sono state rispettivamente di 50, e 52 e 60 unità. Cioè:

Mese	Passo temporale	Vendite
Settembre	t-2	50
Ottobre	t-1	52
Novembre	t (adesso)	60
Dicembre	t+1	

# Esempio

Le vendite dello shampoo al tempo  $t+1$  sono date da

- l'incremento medio negli  $n$  mesi precedenti,
- applicato sulle vendite al tempo  $t$ .

Mese	Passo temporale	Vendite
Settembre	$t-2$	50
Ottobre	$t-1$	52
Novembre	$t$ (adesso)	60
Dicembre	$t+1$	

# Esempio

Le vendite dello shampoo al tempo  $t+1$  sono date da

- l'incremento medio negli  $n$  mesi precedenti,
- applicato sulle vendite al tempo  $t$ .

Mese	Passo temporale	Vendite
Settembre	$t-2$	50
Ottobre	$t-1$	52
Novembre	$t$ (adesso)	60
Dicembre	$t+1$	$(2+8)/2 + 60 = 65$

# Spunto per l'implementazione del modello

```
class Model(object):  
  
    def fit(self, data):  
        pass  
  
    def predict(self):  
        pass  
  
class IncrementModel(Model):  
  
    def fit(self, data):  
        raise NotImplementedError('Questo modello non prevede un fit')  
  
    def predict(self, prev_months):  
        # Qui ci va il codice per far funzionare la predizione.  
        # Nota: "prev_months" deve contenere i dati degli "n" mesi precedenti  
        pass
```

# Esercizio

Finite l'oggetto **CSVFile** come da esercizi delle lezioni precedenti. In particolare, assicuratevi che possa tornare solo un range di linee.

Poi, create il vostro primo modello  
come abbozzato nella slide precedente!

*Ricordatevi di chiedervi che input possono arrivare e come controllare che siano corretti, e che se volete potete anche scrivere un paio di test*