

# *Analisi del Mercato del Fotovoltaico*

Strumenti Statistici per l'Analisi di Dati Aziendali

Benetti Matteo, Coppo Rachele, Fanesi Alessandro

A.A. 2022/2023



# Presentazione dell'azienda

L'azienda per la quale è stata condotta questa analisi di business ha sede in provincia di Vicenza e si occupa della vendita di impianti fotovoltaici, idroelettrici e di trattamento per biomasse e rifiuti per industrie o privati.



# Sommario

- 1 Introduzione
- 2 Prima domanda di ricerca
- 3 Seconda domanda di ricerca
- 4 Terza domanda di ricerca



# Sommario

- 1 **Introduzione**
  - Il problema di business
  - Mercato del fotovoltaico
- 2 Prima domanda di ricerca
- 3 Seconda domanda di ricerca
- 4 Terza domanda di ricerca



# Domande di ricerca

## Domanda 1

Come si posiziona l'Italia nel mercato del fotovoltaico rispetto ad altri paesi europei?

## Domanda 2

- 2.1 Quali sono, per l'azienda, i fattori che incidono sul prezzo di vendita dei pannelli fotovoltaici?
- 2.2 Nel periodo che va dal 2020 al 2022 come si è sviluppato il trend delle vendite dell'azienda?

## Domanda 3

Le politiche di incentivazione proposte dal governo italiano nel 2021 hanno aumentato la propensione dei clienti dell'azienda ad investire di più (o di meno) nel fotovoltaico?

# Descrizione del mercato del fotovoltaico

Principali fattori che incidono sul mercato del fotovoltaico:

- Costo dei pannelli fotovoltaici;
- Politiche ed incentivi governativi;
- Accettazione sociale;
- Avanzamenti tecnologici;
- Condizioni climatiche.



# Politiche ed incentivi governativi

Nella fase di sviluppo del prodotto sono state necessarie forti politiche di incentivazione:

- Polizze assicurative di supporto;
- Feed-in-tariffs (FIT);
- Sussidi;
- Premi.



# Sommario

- 1 Introduzione
- 2 **Prima domanda di ricerca**
  - Fattori Climatici
  - Installazioni
  - Conclusioni
- 3 Seconda domanda di ricerca
- 4 Terza domanda di ricerca



## Domanda 1

Come si posiziona l'Italia in termini di installazioni di impianti fotovoltaici rispetto ad altri paesi europei?



# Dataset

- Fonte: Climate Data Store (Copernicus)<sup>1</sup>;
- Periodo di riferimento: dal 1992 al 2021;
- Frequenza rilevazione: giornaliera;
- 38 paesi selezionati;
- Variabili climatiche:
  - ▶ *prec*: precipitazioni totali (*m* per unità d'area);
  - ▶ *press*: pressione atmosferica al livello del mare (*hPa*);
  - ▶ *radiation*: radiazioni solari (*W/m<sup>2</sup>*);
  - ▶ *temp*: temperatura dell'aria a due metri d'altezza (*K*);
  - ▶ *wind\_10*: velocità del vento a 10 metri di altezza (*m/s*);
  - ▶ *wind\_100*: velocità del vento a 100 metri di altezza (*m/s*).

---

<sup>1</sup><https://cds.climate.copernicus.eu>



# Analisi preliminari

- Per ogni variabile:
  - ▶ Eliminazione degli stati con valori mancanti;
  - ▶ Aggregazione dei dati tramite la media.
- Creazione del dataset contenente **36 osservazioni** e **6 variabili**;
- Le variabili sono state standardizzate.



# Analisi di segmentazione degli stati

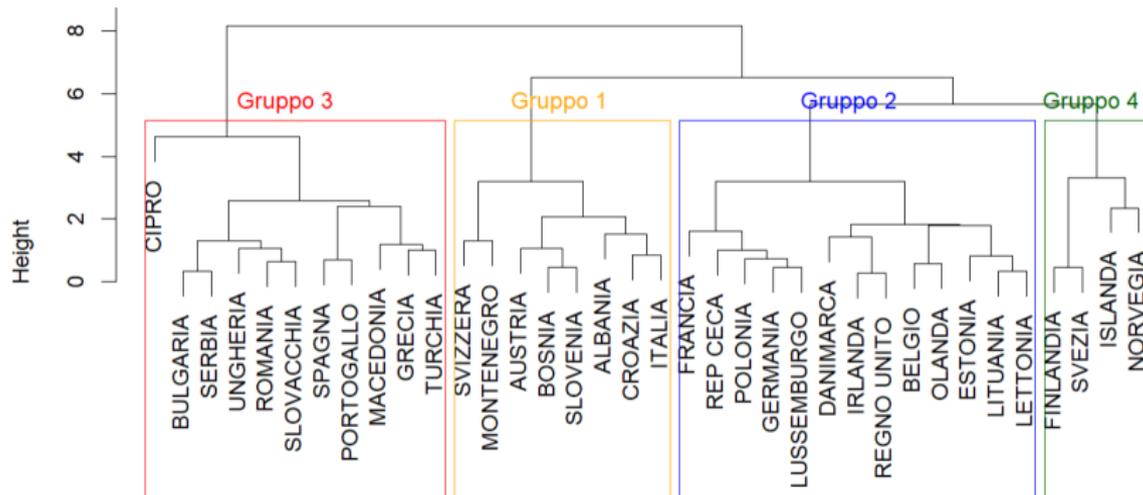
## Metodologie

- 1 Calcolo della distanza euclidea;
- 2 Uso del clustering gerarchico con legame singolo, completo e medio;
- 3 Uso del clustering non gerarchico con il metodo delle K-medie.



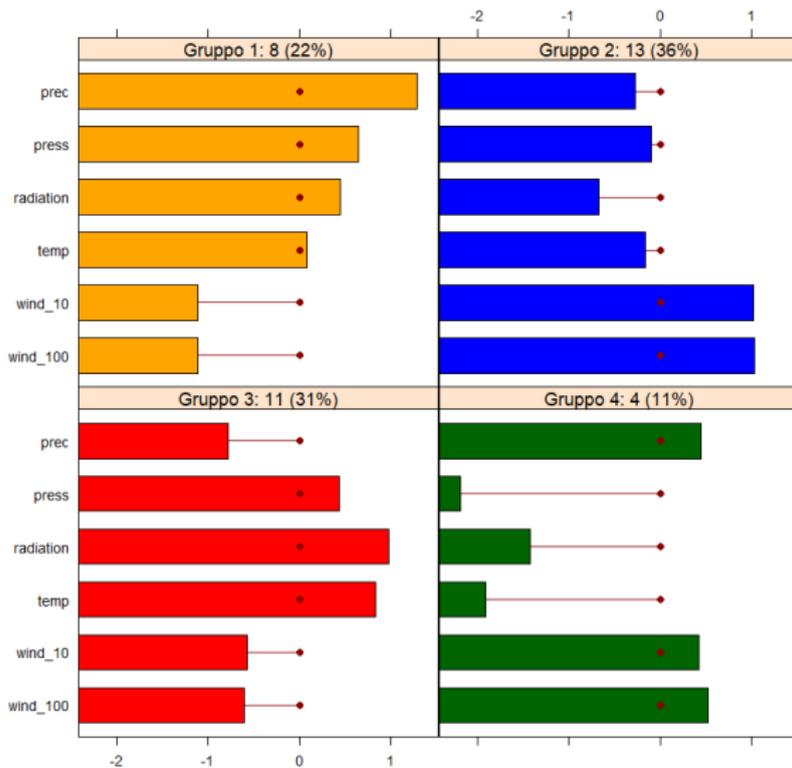
# Analisi di segmentazione degli stati

## Dendrogramma - Legame completo



# Analisi di segmentazione degli stati

## Barchart



# Analisi di segmentazione degli stati

## Commenti

Vengono individuati 4 gruppi:

- Gruppo 1: Albania, Austria, Bosnia, Svizzera, Croazia, Italia, Montenegro e Slovenia.  
Hanno caratteristiche meteorologiche adeguate all'installazione degli impianti fotovoltaici;
- Gruppo 2: Belgio, Repubblica Ceca, Germania, Danimarca, Estonia, Francia, Irlanda, Lituania, Lussemburgo, Lettonia, Olanda, Polonia e Regno Unito.  
Le condizioni climatiche non sono ottimali;



# Analisi di segmentazione degli stati

## Commenti

- Gruppo 3: Bulgaria, Cipro, Grecia, Spagna, Ungheria, Macedonia, Portogallo, Romania, Serbia, Slovacchia e Turchia.  
Condizioni climatiche migliori;
- Gruppo 4: Norvegia, Finlandia, Islanda e Svezia.  
Hanno il clima peggiore per l'installazione del fotovoltaico.



# Dataset PVPS

- Fonte: rapporto di IEA-PVPS<sup>2</sup>del 2022;
- Stati europei selezionati: Italia, Francia, Germania, Spagna, Belgio, Austria, Danimarca, Finlandia, Olanda, Portogallo, Norvegia, Svezia e Svizzera;
- Dati: serie storica annuale dal 1992 al 2021;
- Variabile rilevata: capacità produttiva nominale del fotovoltaico ( $MW_p$ ).

---

<sup>2</sup><https://iea-pvps.org/trends-reports/trends-2022/>



# Dataset popolazione

- Le installazioni sono state standardizzate con i dati delle popolazioni degli stati per renderle confrontabili;
- I dati delle popolazioni<sup>3</sup> sono in forma di serie storica annuale nello stesso periodo di riferimento delle installazioni (1992-2021);
- Il dataset costruito contiene **30 osservazioni**, una per ciascun anno, relative ai **13 stati** considerati.

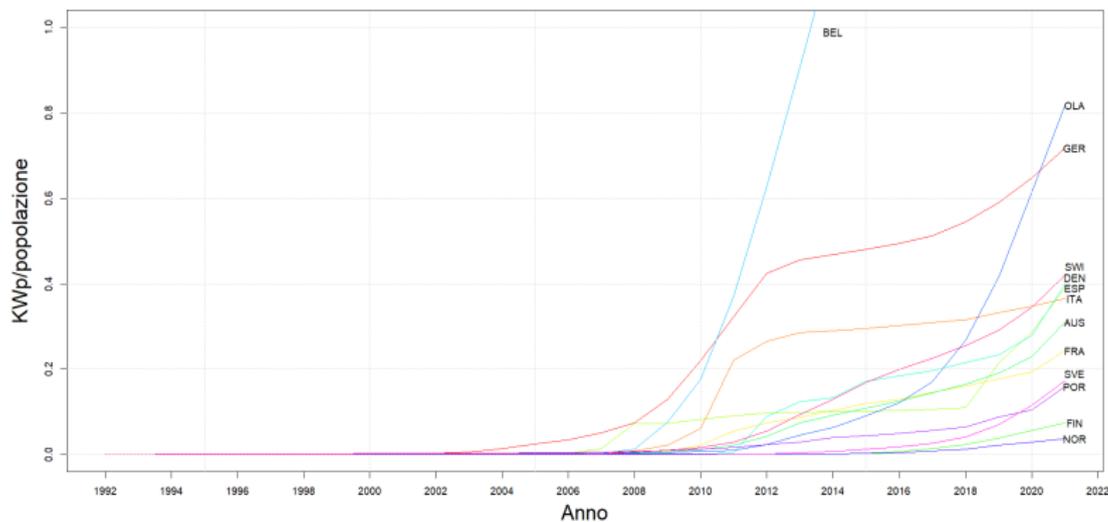
---

<sup>3</sup><https://databank.worldbank.org>



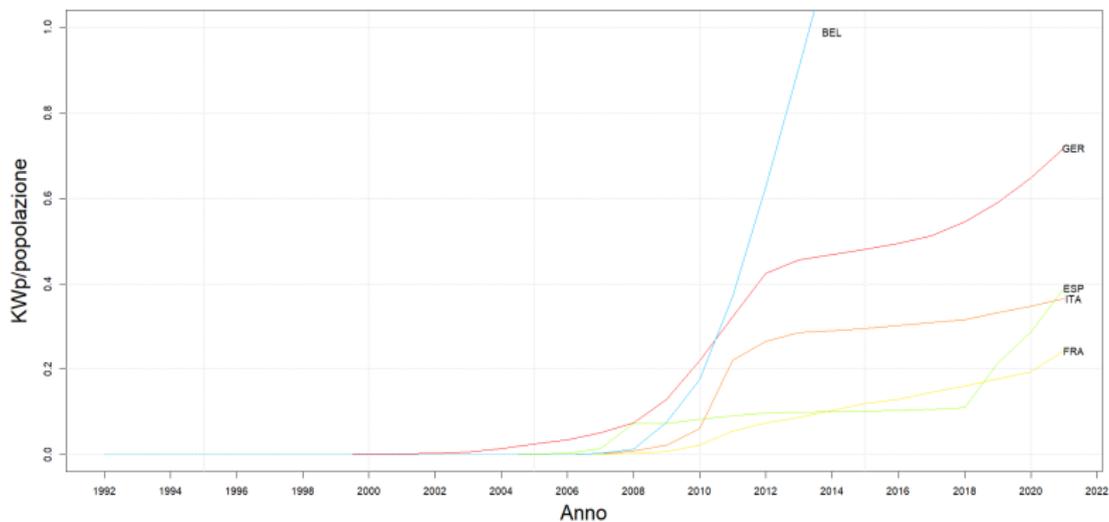
# Capacità produttiva cumulata delle installazioni standardizzata per la popolazione

Tutti i paesi



# Capacità produttiva cumulata delle installazioni standardizzata per la popolazione

Paesi di interesse



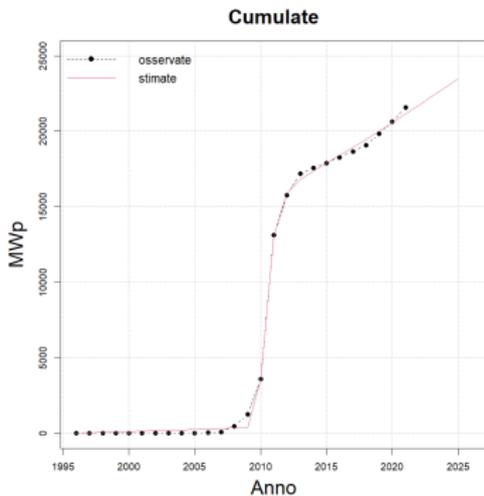
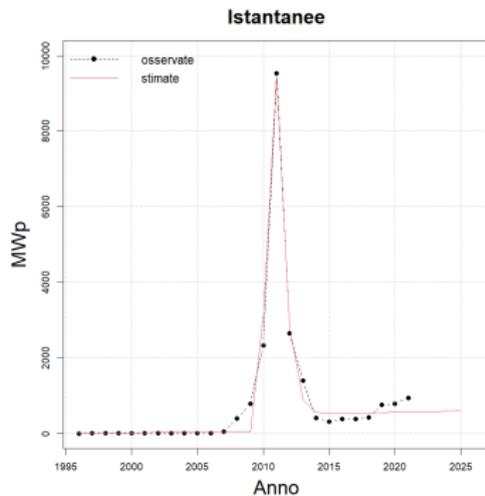
# Modelli di Bass

- Per ognuno dei 5 paesi selezionati viene adattato un modello di Bass standard e si nota che l'adattamento non è soddisfacente;
- Si adatta un modello di Bass generalizzato, inserendo shock esponenziale o rettangolare;
- Si riporta solamente l'output del modello migliore.



# Modello di Bass generalizzato

## Italia - Shock esponenziale



# Modello di Bass generalizzato

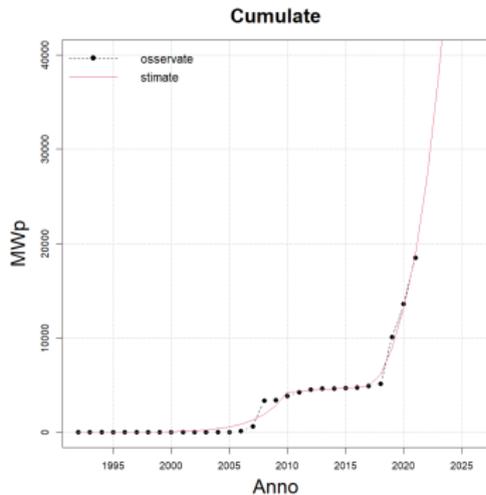
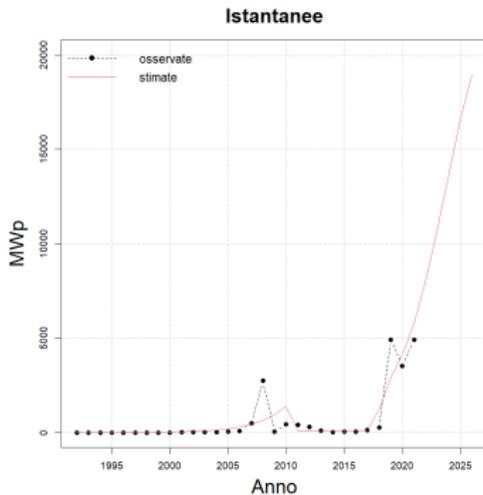
Italia - Shock esponenziale

	Estimate	Std.Error	$p$ -value
m	74383	262262	0.78
p	0.0003	0.001	0.77
q	0.0367	0.0475	0.45
a1	14.648	0.0715	<b>&lt;0.0001</b>
b1	-1.8715	0.2931	<b>&lt;0.0001</b>
c1	148.31	128.49	0.26



# Modello di Bass generalizzato

Spagna - Shock rettangolare



# Modello di Bass generalizzato

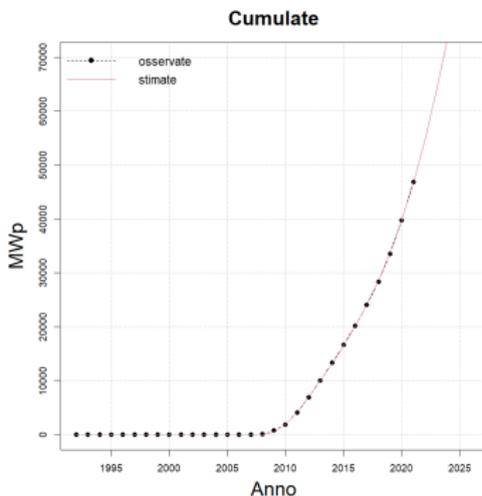
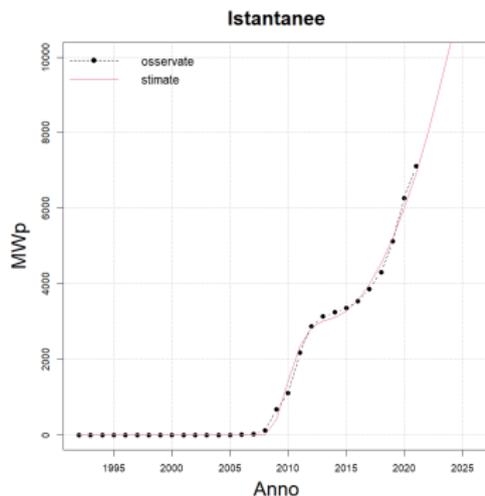
Spagna - Shock rettangolare

	Estimate	Std.Error	<i>p</i> -value
m	202056	695927	0.77
p	0.000005	0.00001	0.63
q	0.3976	0.0764	< <b>0.0001</b>
a1	19.016	0.3884	< <b>0.0001</b>
b1	26.436	0.3048	< <b>0.0001</b>
c1	-0.9505	0.0556	< <b>0.0001</b>



# Modello di Bass generalizzato

## Belgio - Shock esponenziale



# Modello di Bass generalizzato

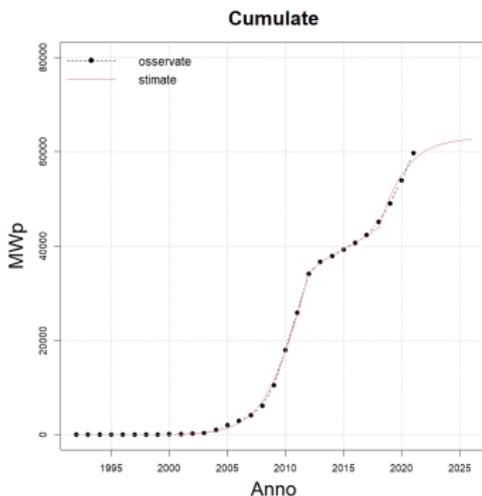
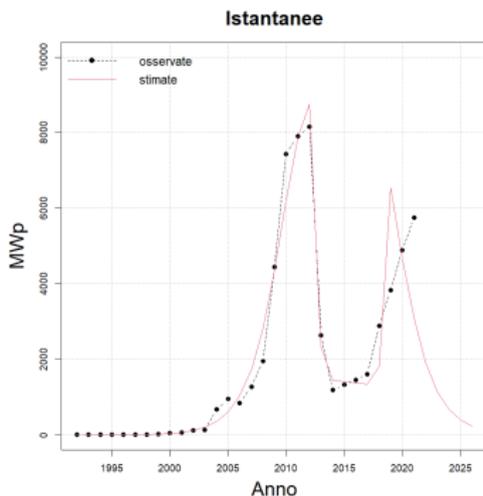
Belgio - Shock esponenziale

	Estimate	Std.Error	$p$ -value
m	550849	516153	0.29
p	0.000001	0.00002	0.42
q	0.1725	0.0123	< <b>0.0001</b>
a1	17.208	0.3693	< <b>0.0001</b>
b1	-0.6108	0.0436	< <b>0.0001</b>
c1	14.701	3.1612	< <b>0.0001</b>



# Modello di Bass generalizzato

## Germania - Shock rettangolare



# Modello di Bass generalizzato

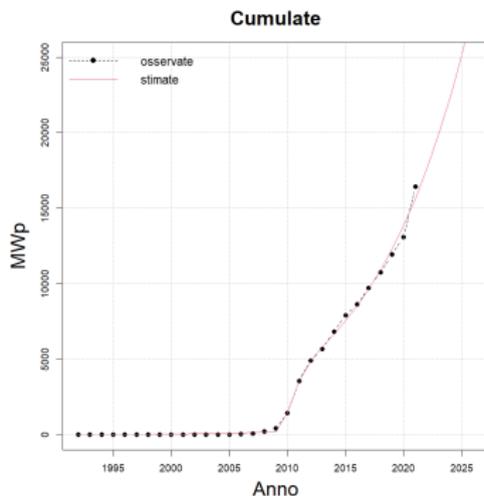
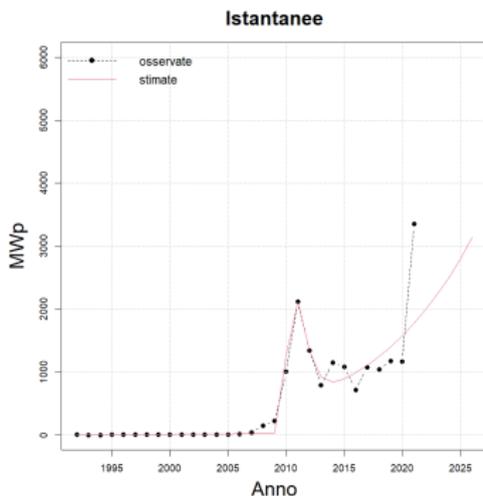
Germania - Shock rettangolare

	Estimate	Std.Error	<i>p</i> -value
m	62873	992.75	< <b>0.0001</b>
p	0.000005	0.000001	< <b>0.0001</b>
q	0.5623	0.0185	< <b>0.0001</b>
a1	21.122	0.1176	< <b>0.0001</b>
b1	26.917	0.1352	< <b>0.0001</b>
c1	-0.8314	0.0253	< <b>0.0001</b>



# Modello di Bass generalizzato

## Francia - Shock esponenziale



# Modello di Bass generalizzato

Francia - Shock esponenziale

	Estimate	Std.Error	$p$ -value
m	534210	6082980	0.93
p	0.000005	0.00006	0.93
q	0.1234	0.0313	< <b>0.001</b>
a1	18.274	0.3234	< <b>0.0001</b>
b1	-1.3683	0.3806	< <b>0.01</b>
c1	31	10.278	< <b>0.01</b>



# Conclusioni

Dagli output dei modelli si evince:

- La componente imitativa ( $p$ ) dei modelli GBM è sempre non significativa, tranne in Germania;
- La crescita del mercato è avvenuta in corrispondenza di periodi in cui venivano emanate politiche di incentivazione;
- Belgio e Germania sono i paesi in cui il mercato del fotovoltaico funziona di più;
- In Italia la crescita del fotovoltaico sembra essersi bloccata;
- Le installazioni in Spagna e Francia negli ultimi anni stanno crescendo.



# Sommario

- 1 Introduzione
- 2 Prima domanda di ricerca
- 3 Seconda domanda di ricerca**
  - Domanda 2.1: Analisi preliminari
  - Domanda 2.1: Modelli previsivi
  - Domanda 2.2: Modello di Bass
- 4 Terza domanda di ricerca



# Obiettivi

## Domanda 2

2.1 Quali sono, per l'azienda, i fattori che incidono sul prezzo di vendita dei pannelli fotovoltaici?



# Dataset azienda

- Fonte: si è preso contatto con un impiegato dell'ufficio vendite dell'azienda che ha messo a disposizione dei dati relativi alle vendite di impianti fotovoltaici a industrie e privati.
- Il dataset è composto da **226 osservazioni** e **9 variabili**:
  - ▶ *date*: giorno della vendita;
  - ▶ *amount*: ammontare della vendita in €;
  - ▶ *power*: potenza in KWp dell'impianto;
  - ▶ *battery*: variabile indicatrice della presenza della batteria nell'impianto (0 se assente, 1 se presente);
  - ▶ *storage*: capacità della batteria installata (KWh);



# Dataset azienda

- ▶ *industrial*: variabile indicatrice della tipologia di impianto venduto (0 se residenziale, 1 se industriale);
- ▶ *wallbox*: variabile indicatrice della presenza di una colonnina di ricarica per l'automobile (0 se assente, 1 se presente);
- ▶ *discount110*: variabile indicatrice dell'utilizzo dell'ecobonus 110% (0 se assente, 1 se presente);
- ▶ *discount50*: variabile indicatrice dell'utilizzo della detrazione fiscale al 50% (0 se assente, 1 se presente).



# Pulizia dei dati e analisi esplorativa

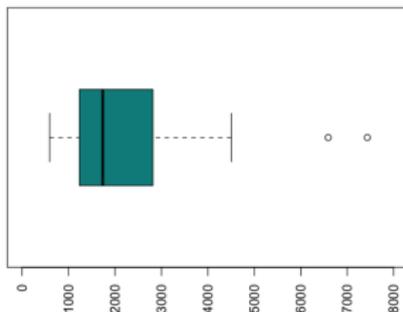
## Variabile risposta

La variabile risposta scelta per condurre l'analisi è stata chiamata *CUP* (Cost for Unit Power)

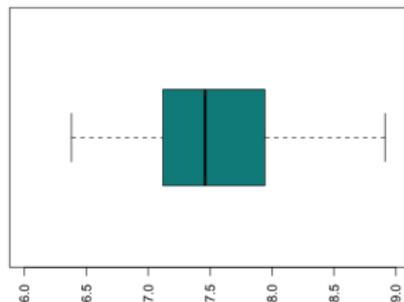
$$CUP = \frac{\text{amount}}{\text{power}},$$

di cui si considera la trasformata logaritmica  $\log(CUP)$ .

Distribuzione di CUP



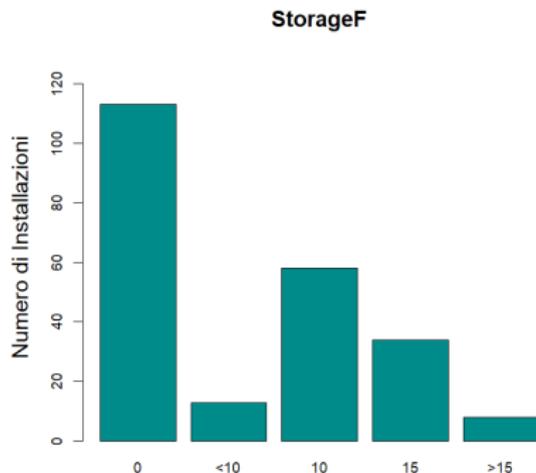
Distribuzione di  $\log(CUP)$



# Pulizia dei dati e analisi esplorativa

## Variabile *storage*

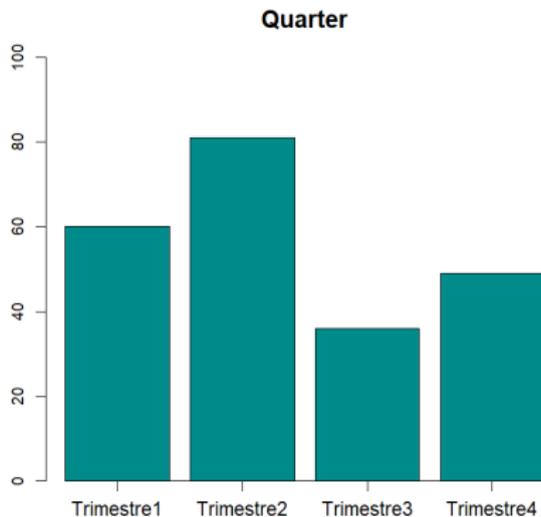
- I valori mancati nella variabile *storage* sono stati imputati con il valore 0;
- La variabile *storage* è stata raggruppata in classi;
- La nuova variabile *storageF* è composta da **5 modalità**: "0", "<10", "10", "15" e ">15".



# Pulizia dei dati e analisi esplorativa

## Variabile *quarter*

- Si è deciso di valutare se ci fosse un effetto di stagionalità trimestrale;
- È stata creata la variabile *quarter* fattore con **4 modalità**:  
"Trimestre1", "Trimestre2", "Trimestre3", "Trimestre4".



# Metodologia e metrica di errore

- Metrica di errore (Mean Squared Error):

$$MSE = \frac{1}{n} \sum_{i=1}^n (\log(CUP_i) - \log(\widehat{CUP}_i))^2;$$

- Selezione del modello e valutazione dell'errore di previsione:  
Convalida incrociata con 5 gruppi.
- Variabile esplicative utilizzate: *industrial*, *wallbox*, *storageF*, *quarter*.



# Selezione del miglior modello

Errori di previsione calcolati in convalida incrociata

Modello	MSE CV	Sd CV
Albero	0.0522	0.0093
Lasso	0.0508	0.0043
<b>Mars</b>	<b>0.0421</b>	<b>0.0114</b>
Random Forest	0.0495	0.0080
Ridge	0.0551	0.0110
Gbm	0.0551	0.0110



# Risultati

Tabella delle variabili selezionate dal miglior modello (Mars)

Predittore	Livello	Coefficiente	Sd
Intercetta	NA	7.9742	0.0253
StorageF	"0"	-0.7111	0.04
StorageF	"<10"	-0.2549	0.0771
Industrial	"1"	-0.3572	0.0490



# Obiettivi

## Domanda 2

2.2 Nel periodo che va dal 2020 al 2022 come si è sviluppato il trend delle vendite dell'azienda?



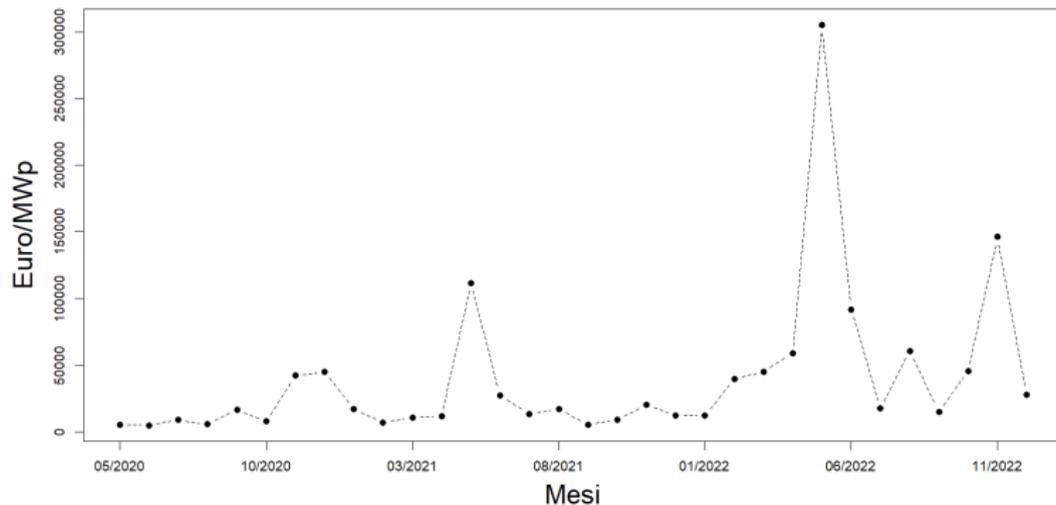
# Preparazione dei dati

- Per rispondere alla domanda 2.2 si è scelto di valutare l'andamento delle vendite tramite la variabile *amount*;
- Si considera la media mensile per i 3 anni;
- Si nota che nel mese di aprile 2020 non sono state rilevate vendite;
- Nei mesi di gennaio, febbraio e marzo sono presenti solo 5 osservazioni che si è deciso di eliminare.



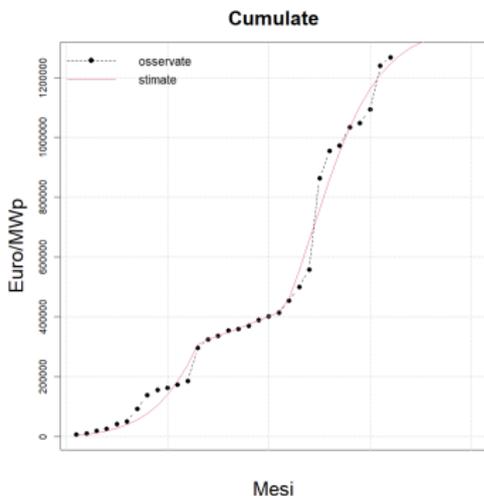
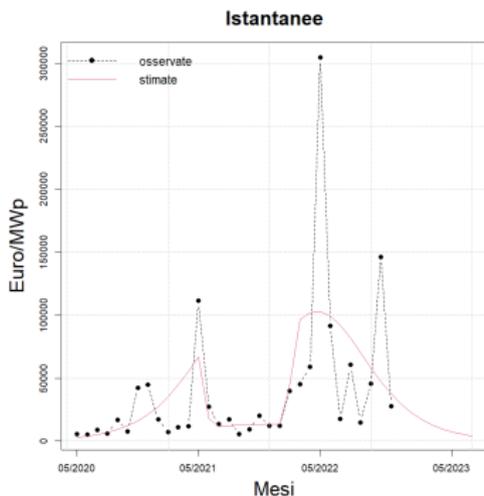
# Serie storica delle vendite

Media mensile di *amount*



# Modello di Bass generalizzato

## Shock rettangolare



# Modello di Bass generalizzato

## Shock rettangolare

	Estimate	Std.Error	<i>p</i> -value
m	1368780	79464	< <b>0.0001</b>
p	0.002	0.0006	< <b>0.01</b>
q	0.2986	0.0426	< <b>0.0001</b>
a1	13.099	0.7224	< <b>0.0001</b>
b1	21.608	0.6157	< <b>0.0001</b>
c1	-0.8437	0.0878	< <b>0.0001</b>



# Sommario

- 1 Introduzione
- 2 Prima domanda di ricerca
- 3 Seconda domanda di ricerca
- 4 Terza domanda di ricerca**
  - Politiche di incentivazione
  - Analisi preliminari
  - Modelli previsivi
  - Conclusioni



# Terza domanda di ricerca

## Domanda 3

Le politiche di incentivazione proposte dal governo italiano nel 2021 hanno aumentato la propensione dei clienti dell'azienda ad investire di più (o di meno) nel fotovoltaico?



# Agevolazioni fiscali

- Agevolazioni fiscali:
  - ▶ ECOBONUS 110%,
  - ▶ Detrazione fiscale al 50%;
- Registrate osservazioni solo nel 2021.



# Dataset azienda 2021

- Dataset: **72 unità statistiche** e **9 variabili**;
- Viene aggiunta la variabile *discount* (1 se presente una delle due agevolazioni 0 altrimenti);
- Variabile risposta:  $\log(CUP)$ ;
- Variabile esplicative utilizzate: *industrial*, *wallbox*, *storageF*, *quarter* e *discount*.



# Selezione del miglior modello

Errori di previsione calcolati in convalida incrociata

Modello	MSE CV	Sd CV
Albero	0.0577	0.0146
Lasso	0.0638	0.0115
<b>Mars</b>	<b>0.0245</b>	<b>0.028</b>
Random Forest	0.0525	0.018
Ridge	0.0637	0.0101
Gbm	0.0657	0.0365

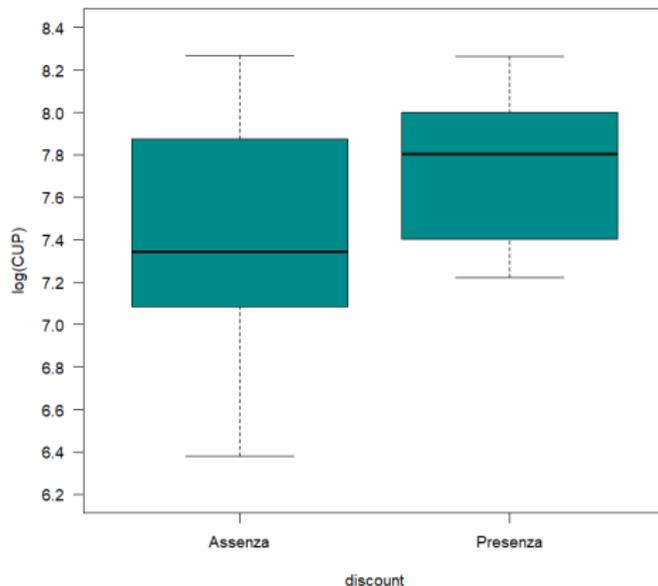


# Risultati dei modelli

- La variabile *discount* non viene selezionata dal miglior modello;
- Nessun modello seleziona la variabile *discount*;
- L'accesso alle agevolazioni fiscali non ha effetto sulla propensione del cliente ad investire di più (o di meno) nel fotovoltaico.



# Distribuzione di $\log(\text{CUP})$ condizionata a discount



Grazie per  
l'attenzione!

