

MODELLO PREDITTIVO

BUDGET VENDITE

Progetto «synAPs»

Qual è il problema che dobbiamo risolvere?

- Il processo di elaborazione del budget delle vendite implica un dispendio di risorse e di tempo data la quantità di informazioni che è difficile da analizzare e controllare.
- Per porre rimedio a questo problema e per elaborare e preventivare il budget delle vendite in modo efficiente e qualitativo, si rende necessario introdurre un modello predittivo con una piattaforma decisionale che consenta agli utenti di sfruttare ed elaborare i dati storici archiviati ed ottenere dei risultati previsionali affidabili ed accurati.



Multinazionale settore chimico
con sedi produttive e logistiche
Worldwide



Vuole dotarsi di uno strumento
che fornisca i dati previsionali di
budget vendite



Budget & Fatturato ultimi 5 anni
molto eterogenei e frammentati
(in accordo con la tipologia del
Business dei mercati in cui opera
e delle organizzazioni di vendita)



Modello previsionale
Efficientare il processo
Facilitare la verifica delle previsioni

**Conoscenza aziendale, conoscenza del dominio,
conoscenza tecnica, conoscenza della metrica**

Formulazione del problema

- La previsione dei dati di budget di vendita è diversa dalle tipiche applicazioni predittive e di Machine Learning, poiché **non** consiste nel replicare attività che gli esseri umani possono svolgere facilmente.
- I dati di budget di vendita sono delle serie temporali, non stazionari e non lineari che sono influenzati da fattori esterni.
- La previsione accurata dei dati di budget è un compito strategico e impegnativo per una gestione ottimale delle risorse, richiede l'utilizzo di un modello il più accurato possibile.
- L'applicazione di metodi predittivi a questi problemi può produrre risultati più utili rispetto ai metodi standard utilizzati in ambito sales e finance.

Convertire il problema aziendale in un problema analitico.

- In questo documento proponiamo un approccio che utilizza e confronta due modelli predittivi:
 - ARIMA
 - Exponential Smoothing

Qual' è la metrica di business?

- Costo di Acquisizione Cliente (CAC)
- La percentuale Marketing del CAC
- Tempo necessario per ripagare il CAC (Payback Time)
- % clienti derivanti dal Marketing
- % clienti influenzati dal Marketing

Il Modello è un approccio appropriato?

- Modelli confrontati:

- ARIMA
- Exponential smoothing

- Modello scelto

- Exponential smoothing

- Motivi della scelta:

- Accuratezza delle previsioni
- Velocità di calcolo

ARIMA $(p,d,q)(P,D,Q)m$

- Un concetto fondamentale nei modelli ARIMA (Autoregressive Integrated Moving Average) è quello della stazionarietà: una serie è stazionaria se le sue proprietà non dipendono dal tempo.
- Un modo per rendere una serie stazionaria è quello di “differenziarla”, cioè calcolare le differenze tra osservazioni consecutive.

Un modello ARIMA viene identificato da tre parametri:

1. p indica l'ordine della parte autoregressiva (AR), che modella la variabile d'interesse usando una combinazione lineare di osservazioni passate;
 2. d indica l'ordine d'integrazione, cioè il numero di differenze operato per rendere la serie stazionaria;
 3. q indica l'ordine della parte a media mobile (MA), che modella la variabile d'interesse usando una combinazione lineare degli errori passati.
- Ci possono essere ulteriori 3 parametri (P, D, Q) che modellano analogamente le serie stagionali.

Exponential smoothing

- Produce previsioni che sono medie pesate delle osservazioni passate, con pesi che decrescono esponenzialmente man mano che si riferiscono a osservazioni più lontane nel tempo
- Approccio classico che è alla base dei modelli indicati con il nome di *state space models*
- Questi modelli sono costituiti da un'equazione che descrive i dati osservati e da equazioni che descrivono le componenti non osservate (o stati) della serie: errore, trend e stagionalità
- Per ogni componente del modello le possibilità sono le seguenti:
 - Errore: modello con errori additivi o modello con errori moltiplicativi
 - Trend: nessun trend, trend additivo o trend additivo smorzato (il trend non è costante nel futuro ma dopo un po' viene attenuato)
 - Stagionalità: nessuna stagionalità, stagionalità additiva o stagionalità moltiplicativa.

Quali dati sono disponibili?

- Budget - TAGETIK
- Fatturazione - SAP
- DDT - Legacy Operation

Possibili estensioni del progetto integrando :

- CRM - Clustering/ Segmentazione Clienti
- Dati di ordinato - SAP: confronto tra i dati predittivi di vendita con l'effettivo ordinato
- Dati Magazzino - Legacy Operation: confronto dei dati di uscita merci e dati predittivi di vendita



Data Set

Budget & Actual

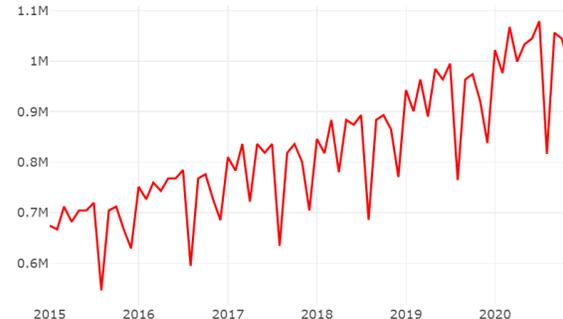
- 2015-2020 (monthly granularity)
- 3 Legal Entity
- 47 Branch
- 13 Business Line

Data Exploration & Evaluation

Data la presenza di serie temporali fortemente irregolari e frammentate si è deciso di considerare soltanto quelle che presentano regolarità stagionale.

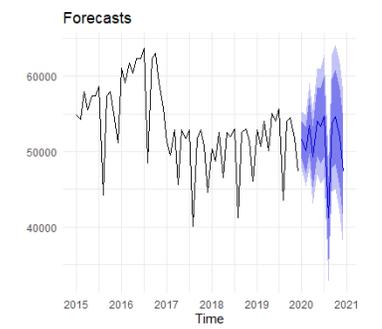
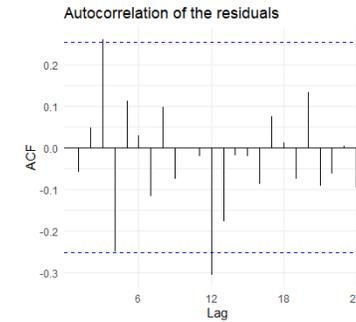


Esempio di serie irregolare

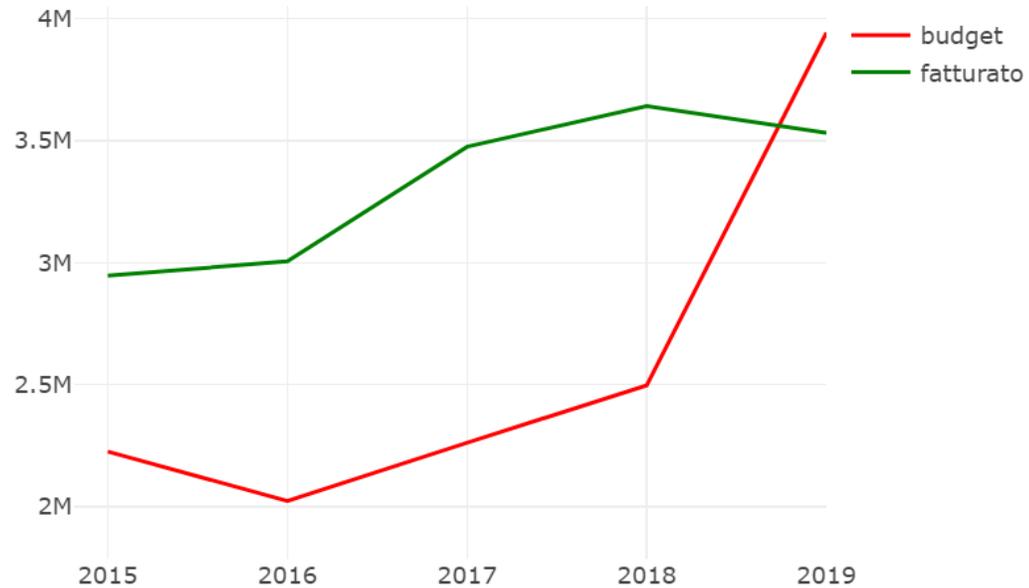


Esempio di serie regolare

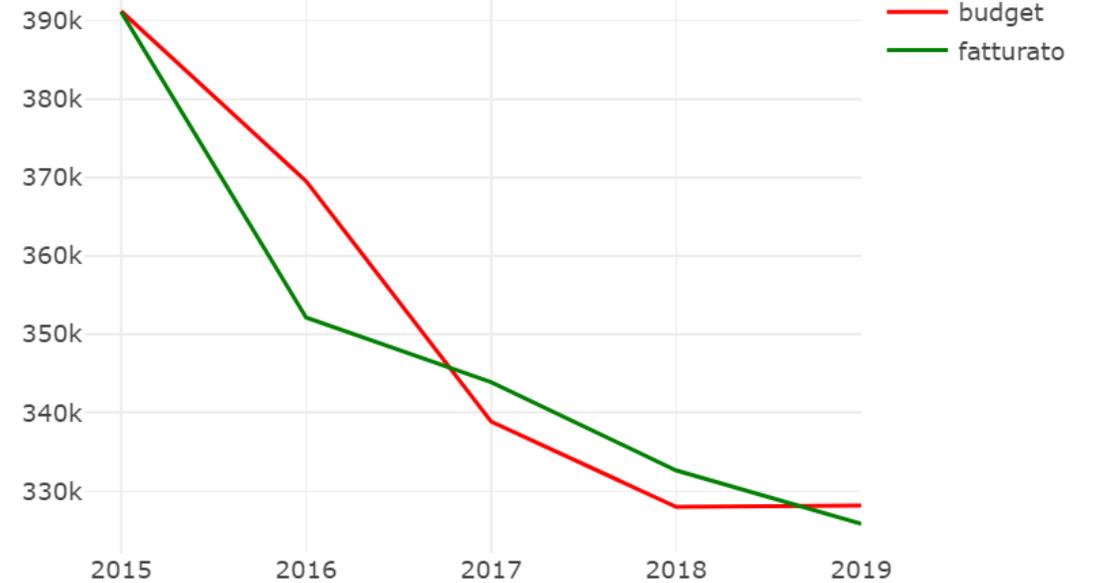
Esempi



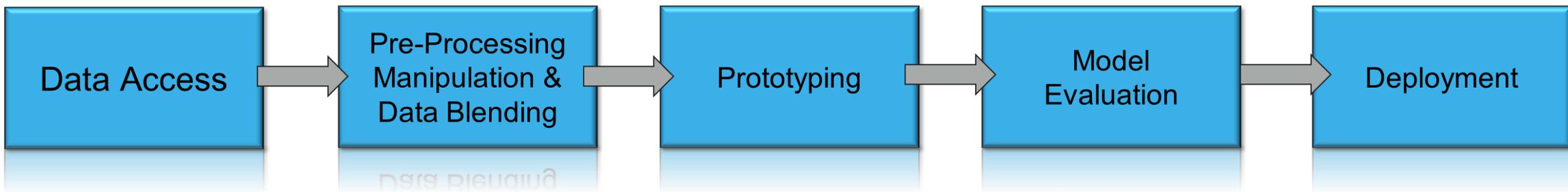
Grafici confronto budget-fatturato anni precedenti



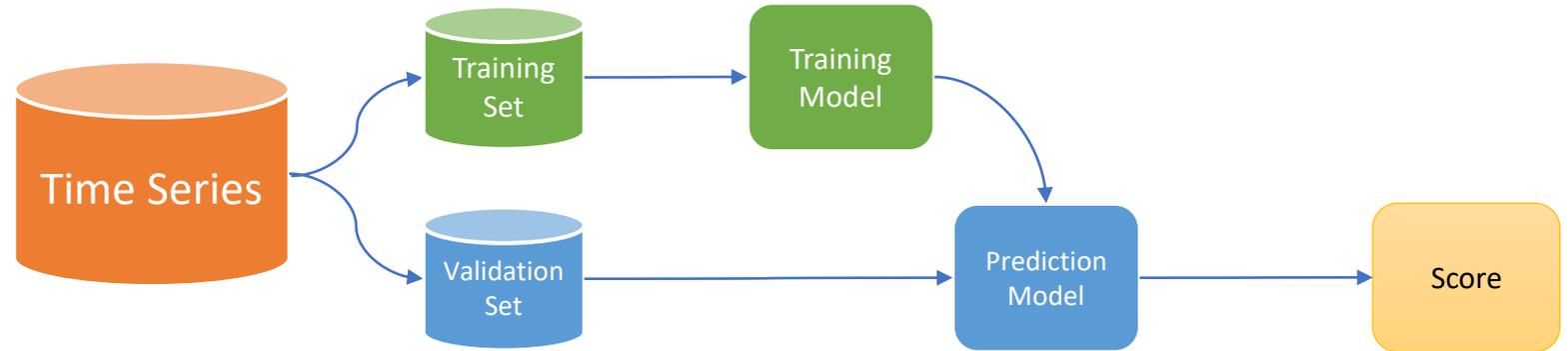
Esempio di una delle serie peggiori



Esempio di una delle serie migliori



Data Mining Process
Machine Learning



Project Phases

- Analisi del modello organizzativo
- Qualità, completezza, granularità dei dati
- Pre-processing
- Manipulation - Data Blending - Binding
- Performance Tuning

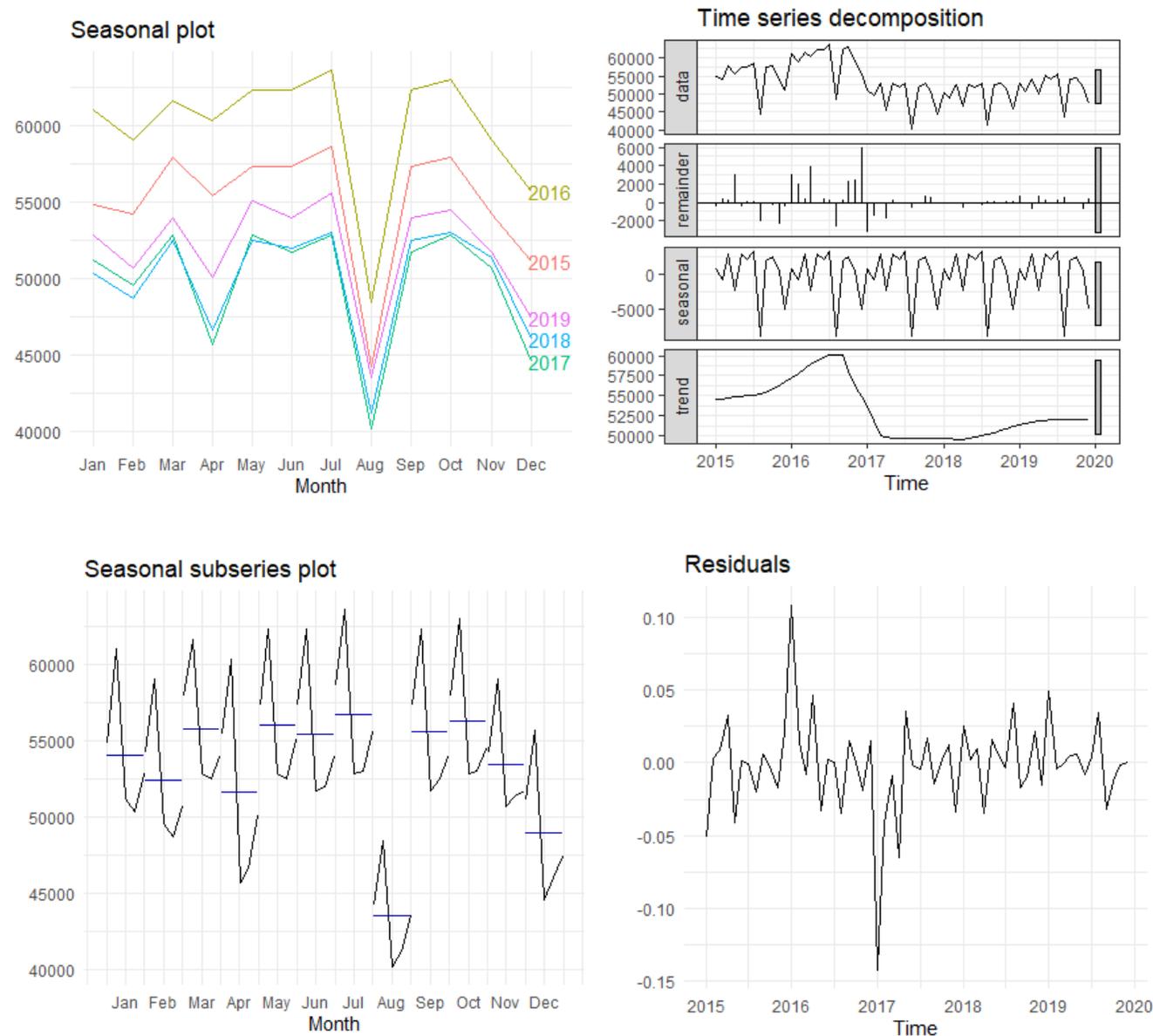
Normalization

Regularizing and Cleaning Time Series

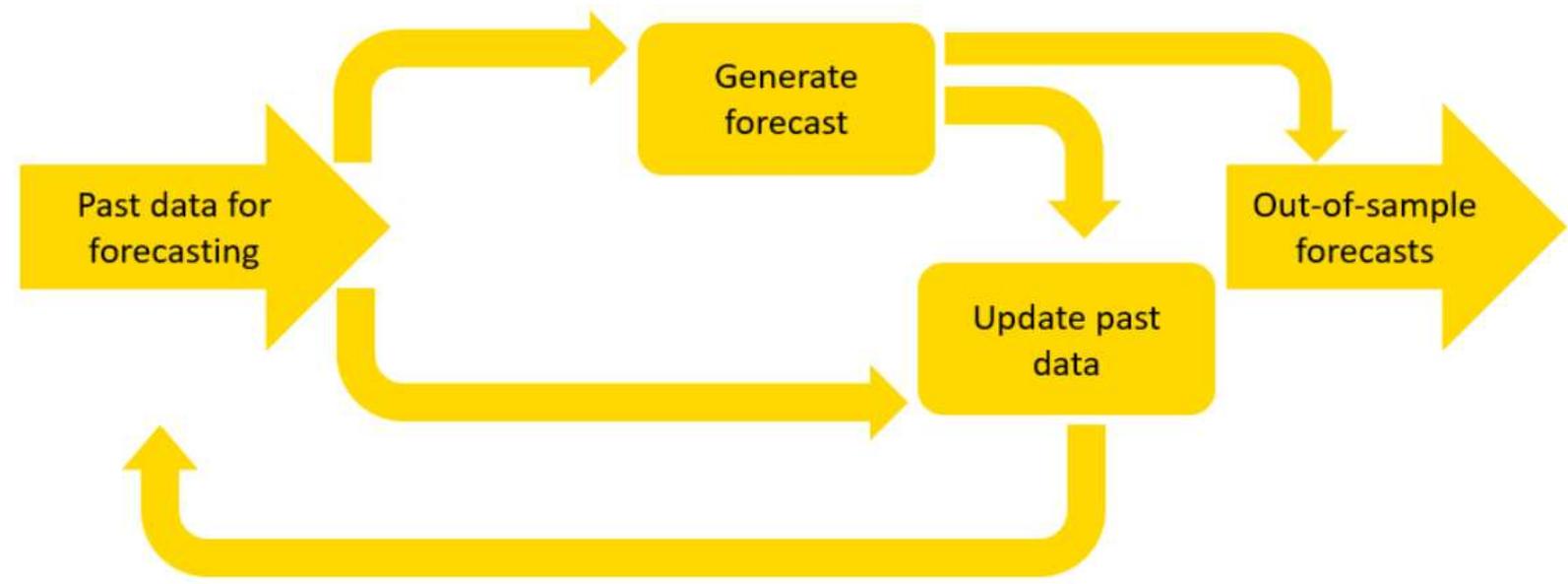
- Complete & Align Time stamp
- Missing Value
- Irregular Pattern

Exploring and Transforming Time Series

- Visual exploring and Transforming
- Classical Decomposition of Time Series

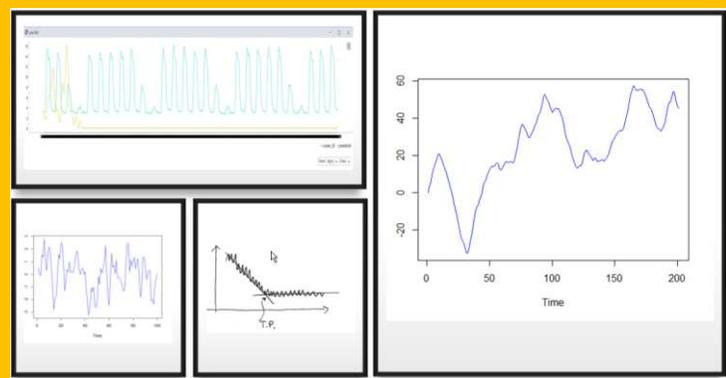


Che tipo di algoritmi eseguiamo?



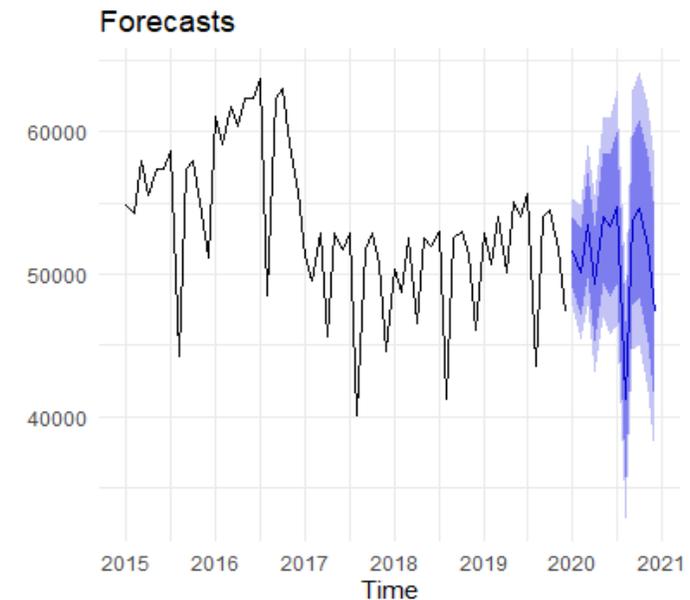
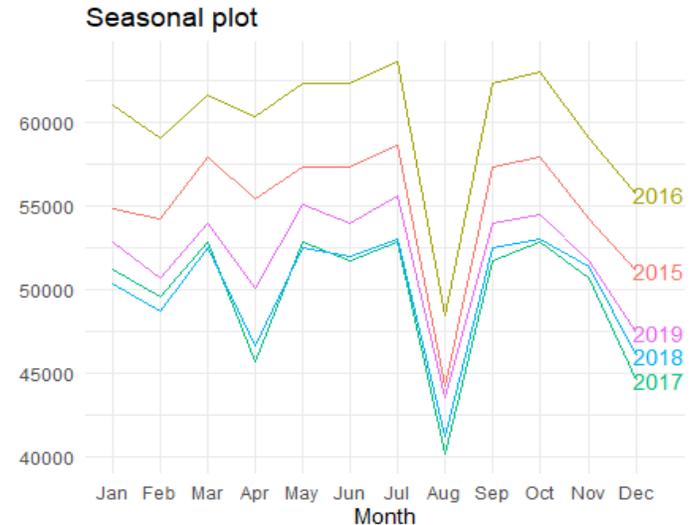
Modelling and Evaluating

- Dynamic Forecasting
- Restoring trend and Seasonalities



Quali sono i nostri obiettivi?

- Il modello predittivo del budget vendite fornisce un importante supporto alla direzione Commerciale e alla direzione Marketing:
 - Immediato riscontro sulla base dell'andamento storico
 - Evidenziare velocemente e con chiarezza le aree che necessitano di particolare attenzione
 - Attraverso il processo di budgeting, favorisce il corretto utilizzo delle risorse aziendali
 - Permette l'analisi degli scostamenti fra previsioni e consuntivi, con diverse granularità

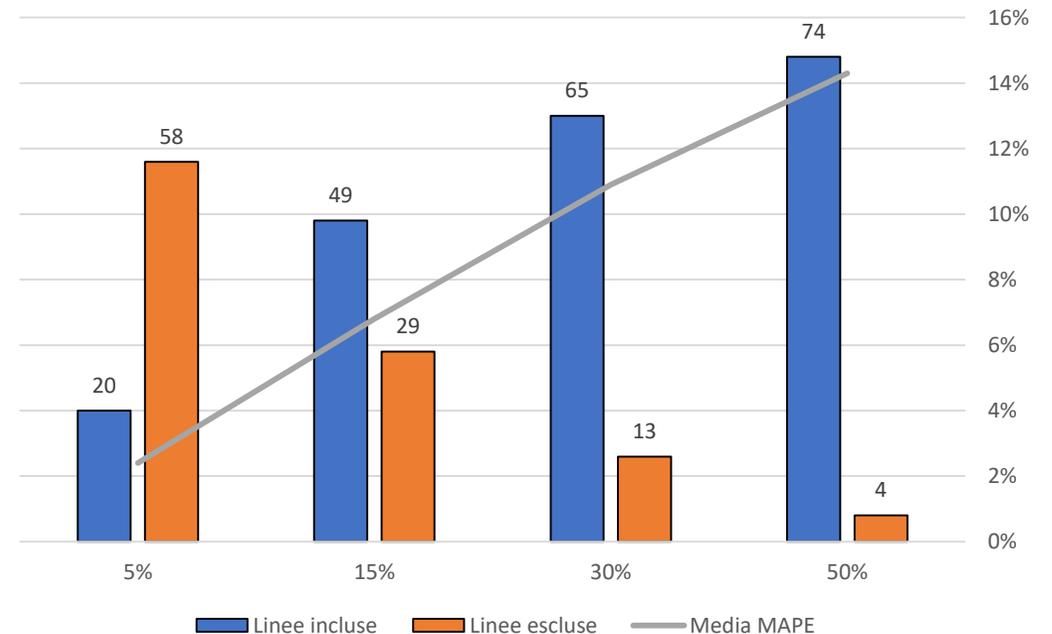


Le serie presentano delle «anomalie» rispetto alle attese, dovute probabilmente a fattori esogeni e comunque degni di ulteriori analisi approfondite

SOGLIA MAPE	Linee incluse	Linee escluse	Media MAPE
5%	20	58	2,40%
15%	49	29	6,80%
30%	65	13	10,90%
50%	74	4	14,30%

- **MAPE: Mean Absolute Percentage Error**

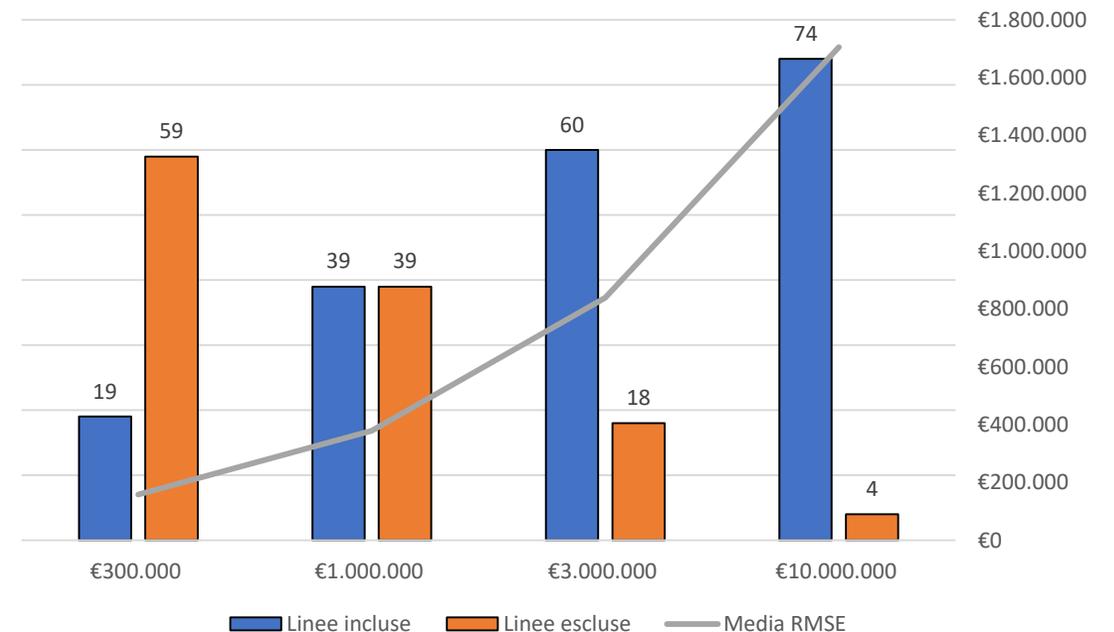
Attendibilità del modello



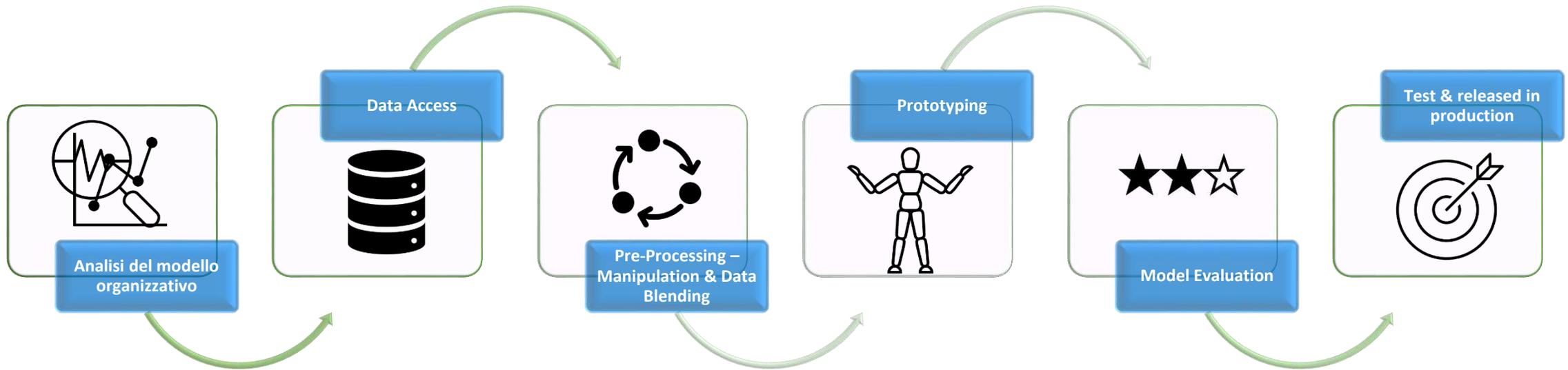
SOGLIA RMSE (€)	Linee incluse	Linee escluse	Media RMSE (€)
300.000 €	19	59	158.480 €
1.000.000 €	39	39	378.656 €
3.000.000 €	60	18	838.295 €
10.000.000 €	74	4	1.706.269 €

- *RMSE: Root Mean Squared Error*

Attendibilità del modello



GANTT



Elapsed time



Team di Lavoro



MARKETING



10%



DATA ENGINEER



10%



DATA SCIENTIST



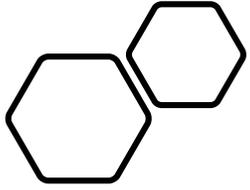
70%



ML ENGINEER

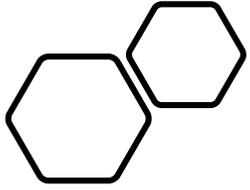


10%



Conclusioni

- Il progetto ha prodotto un modello predittivo per il budget delle vendite, analizzando i dati storici degli ultimi 5 anni.
- Il modello consente di ottenere in brevissimo tempo le previsioni di budget di vendita, potendo fare i confronti con il fatturato e fornendo gli eventuali scostamenti.
- Il modello permette di valutare attraverso dei grafici gli andamenti di budget delle vendite suddivisi per Linea.
- L'azienda dispone di uno strumento che gli consente di ottimizzare in modo rilevante il processo del budget delle vendite.



Abstract

- Le serie temporali finanziarie sono dati non stazionari e non lineari che sono influenzati da fattori esterni. Esistono diversi approcci predittivi performanti come il modello ARIMA e l'Exponential Smoothing. La previsione accurata dei dati di budget è un compito strategico e impegnativo per una gestione ottimale delle risorse, richiede l'utilizzo del modello più accurato.
- Proponiamo un approccio predittivo che utilizza e confronta il modello ARIMA di Machine Learning e il modello Exponential Smoothing. L'applicazione e l'analisi comparativa mostrano che il modello Exponential Smoothing supera il modello ARIMA, principalmente per l'accuratezza delle previsioni e la velocità di calcolo.