

# Aprendizagem de Máquina

## Algoritmos de Previsão Aplicados ao Vestibular da FATEC Mauá: Um Estudo de Séries Temporais

Adriano Barros  
Brendon Gomes  
Elias Barbosa  
Rafael Goncalves

Prof. Bruno Zolotarev

Início

## Objetivo Geral

Analisar dados históricos de inscrições no vestibular da FATEC Mauá para desenvolver modelos de previsão que estimem o número de candidatos nos próximos certames, usando séries temporais e aprendizado de máquina.

## Objetivos Específicos

- . Coleta de dados (Web Scraping - C# (selenium))
- . Tratamento de dados (python e excel)
- . Aplicação dos modelos de previsão Holt-Winters, ARIMA e Prophet aos dados históricos.

# Coleta de Dados

Realizada durante todo o período de divulgação da lista de inscritos Fatec Mauá, desde 2007 até 2025 ( 1º semestre).

Desenvolvida uma automação em C# .NET Framework com Selenium,, que executa web scraping, extraíndo as demandas de vestibulares e exportando os dados em formato CSV.

# Web Scraping



Esse visual não tem suporte para exportação.

Coleta de Dados

Normalização

# Normalização de Dados

## Padronização dos dados

- Excel
- Python (no próprio script)

## Padronização dos nomes

- Períodos ( Noite / Noturno)
- Períodos (Manhã / Tarde - Diurno)
- Nomes dos cursos

## Organização Temporal

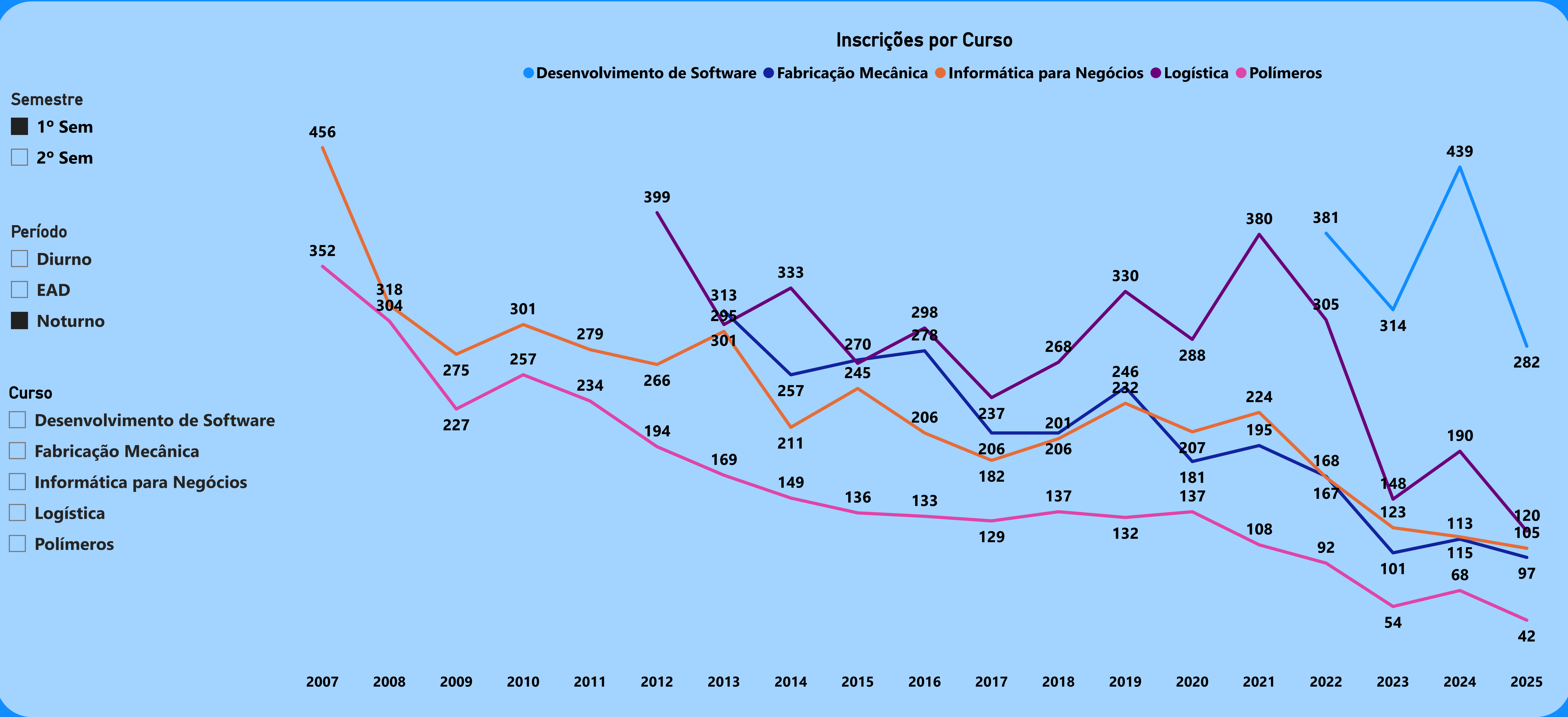
- Por semestres (1º e 2º semestre), devido a particularidade de cada um.

A	B	C	
Ano	Semest	Curso	NomeCurso
2025	1	Polímeros	Polímeros
2024	1	Polímeros	Polímeros
2023	1	Polímeros	Polímeros
2022	1	Polímeros	Polímeros
2021	1	Desenvolvimento de Produtos Plásticos	Polímeros
2021	1	Polímeros	Polímeros
2020	1	Polímeros	Polímeros
2019	1	Desenvolvimento de Produtos Plásticos	Polímeros
2019	1	Polímeros	Polímeros
2018	1	Polímeros	Polímeros
2017	1	Polímeros	Polímeros
2016	1	Polímeros	Polímeros
2015	1	Polímeros	Polímeros
2014	1	Polímeros	Polímeros
2013	1	Polímeros	Polímeros
2012	1	Polímeros	Polímeros
2011	1	Polímeros	Polímeros
2010	1	Polímeros (Produção de Materiais Plásticos)	Polímeros
2009	1	Produção de Materiais Plásticos	Polímeros
2008	1	Produção de Materiais Plásticos	Polímeros
2007	1	Produção de Materiais Plásticos	Polímeros

```
[ ] import pandas as pd
import re

def carregar_e_normalizar_fatec(path="fatec_demanda_data.csv"):
    # Carregar o arquivo
    df = pd.read_csv(path, sep=",")
    # Normalizar nomes das colunas
    df.columns = df.columns.str.strip().str.lower()
    # Normalizar textos
    df["curso"] = df["curso"].str.strip().str.lower()
    df["periodo"] = df["periodo"].str.strip().str.lower()
    df["fatec"] = df["fatec"].str.strip().str.lower()
    # Corrigir nomes de períodos
    df["periodo"] = df["periodo"].replace({
        "noturno": "noite",
        "matutino": "manhã"
    })
    # Padronizar os cursos conforme as regras
    def normalizar_curso(c):
        if "logística" in c:
            return "logística"
        if re.search(r"polímero|plástico", c):
            return "polímeros"
        if "informática" in c:
            return "informática para negócios"
        return c
    df["curso"] = df["curso"].apply(normalizar_curso)
    # Normalizar tipos numéricos
    for col in ["ano", "semestre", "inscritos", "vagas", "demanda"]:
        if col in df.columns:
            df[col] = pd.to_numeric(df[col], errors="coerce")
    return df
```

# Dados Tratados



# Algoritmos Utilizados

## Holt-Winters

Usa médias ponderadas

Peso maior aos dados mais recentes.

Considera:

Tendencia

Sazonalidade

Nível (valor médio atual)

Atualiza os 3 componentes a cada novo dado

Mais usado para séries com padrão sazonal claro e dados mais estáveis

## Prophet

Desenvolvido pelo Facebook,  
Baseado em modelo aditivo que separa a série em:

Tendência (que pode ser linear ou logística),

Sazonalidade,

Feriados e eventos especiais.

Ajusta parâmetros automaticamente, lida bem com sazonalidade.

Robusto a dados faltantes e mudanças repentinas.

## ARIMA

**AutoRegressive (AR)**: valor atual depende dos anteriores.

**Integrated (I)**: diferencia dados para tirar tendências.

**Moving Average (MA)**: usa média dos erros passados

Indicado para séries que precisam eliminar tendências ou sazonalidades.

Flexível, mas precisa escolher bem os parâmetros



# Análise de Tendência

Semestre

1

2

Turma

Noturno

Curso

Desenvolvimento de Software

Fabricação Mecânica

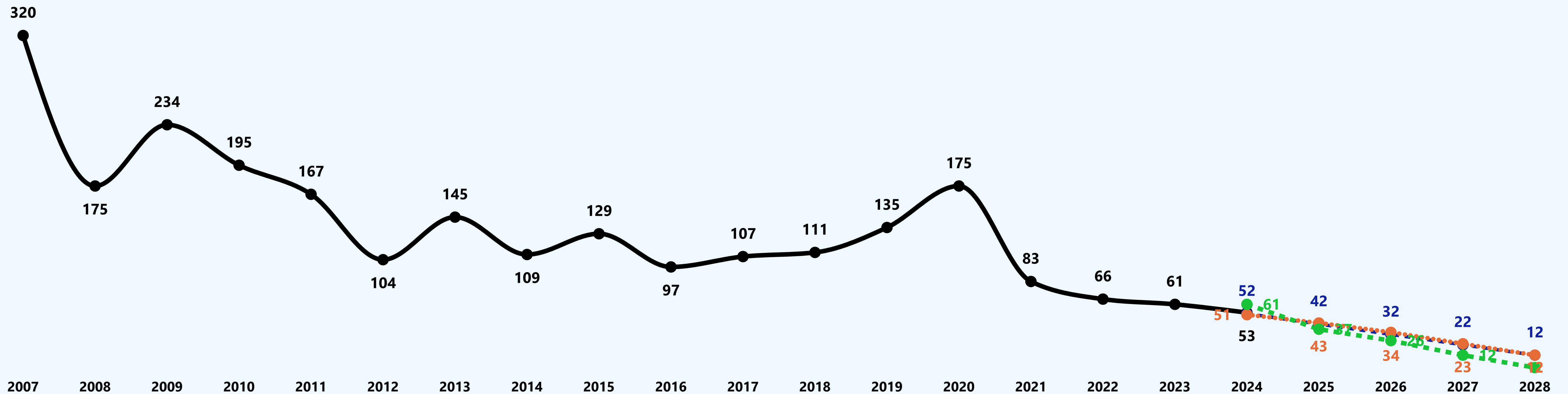
Informática para Negócios

Logística

Polímeros

## Quantidade de Inscritos e Previsões

● Qtd Inscritos ● Previsão Holt ● Previsão Prophet ● Previsão Arima



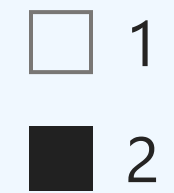
Algoritmos

Holt



# Previsões utilizando Holt-Winters

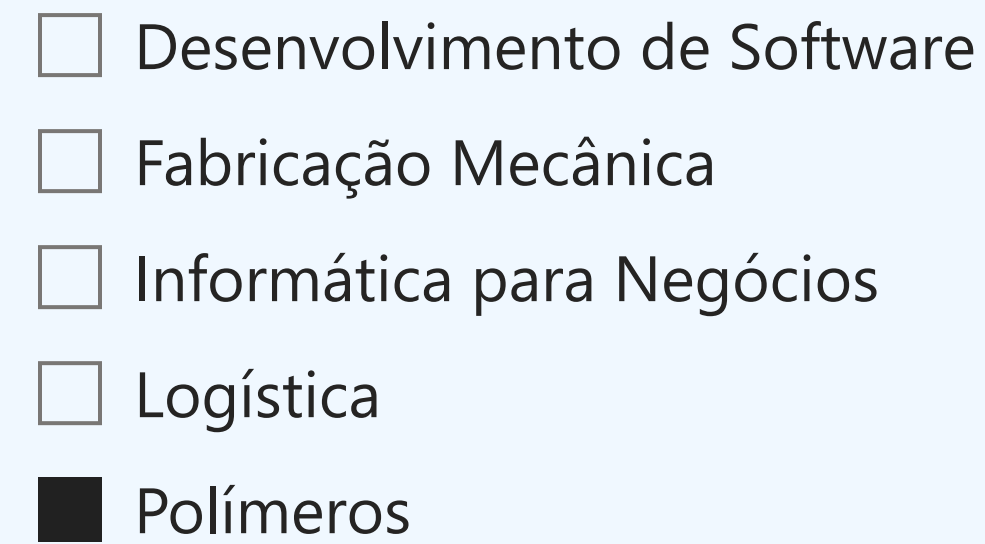
Semestre



Turma

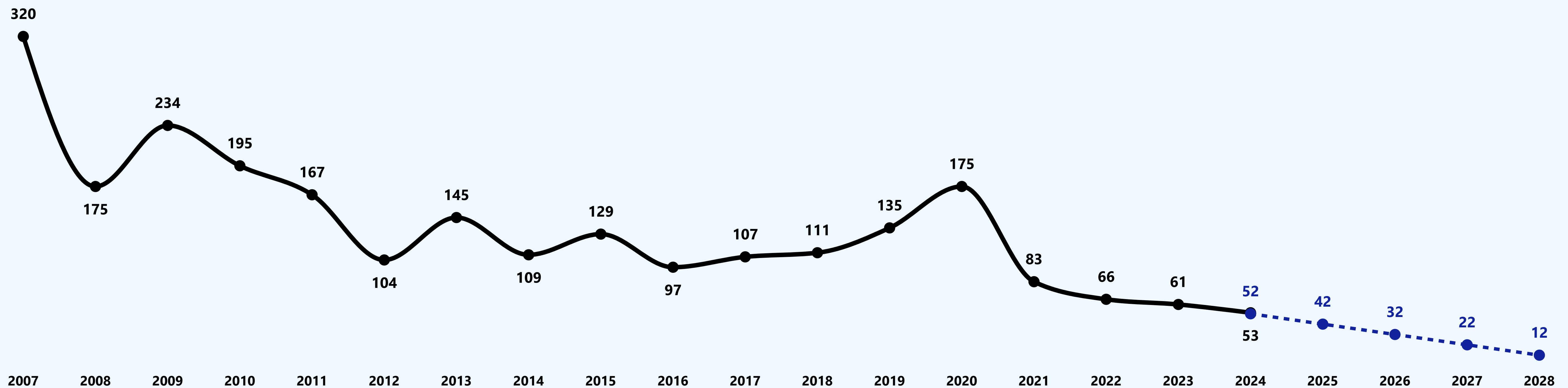


Curso



## Quantidade de Inscritos e Previsões

● Qtd Inscritos ● Previsão Holt



Análise

Prophet

# Previsões utilizando Prophet

Semestre

■ 1  
□ 2

Turma

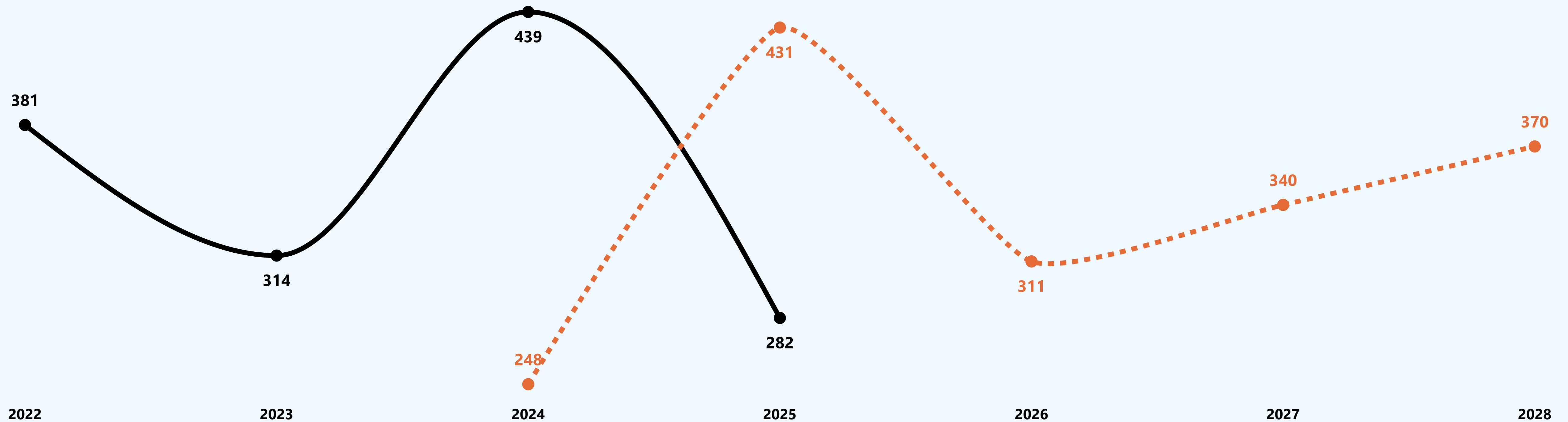
■ Noturno

Curso

■ Desenvolvimento de Software  
□ Fabricação Mecânica  
□ Informática para Negócios  
□ Logística  
□ Polímeros

## Quantidade de Inscritos e Previsões

● Qtd Inscritos ● Previsão Prophet



Holt

ARIMA

# Previsões utilizando ARIMA

Semestre

1

2

Turma

Noturno

Curso

Desenvolvimento de Software

Fabricação Mecânica

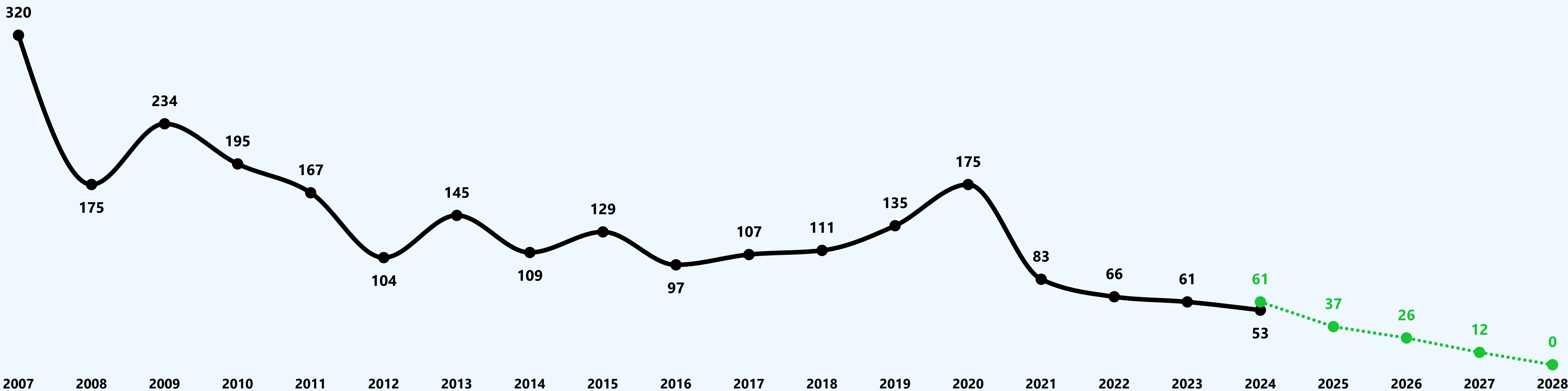
Informática para Negócios

Logística

Polímeros

## Quantidade de Inscritos e Previsões

● Qtd Inscritos ● Previsão Arima



Prophet

Resultados

# Tabela de Comparações

Semestre

1º Sem

2º Sem

Período

Diurno

EAD

Noturno

Ano	Semestre	Curso	Turma	Inscritos	Qtd Prophet	% Prophet	Qtd Holt	% Holt	Qtd Arima	% Arima
2024	1º Sem	Desenvolvimento de Software	Noturno	439	248	-43,5%	247	-43,7%	434	-1,1%
2024	1º Sem	Fabricação Mecânica	Noturno	113	127	12,4%	122	8,0%	57	-49,6%
2024	1º Sem	Informática para Negócios	Noturno	115	135	17,4%	134	16,5%	96	-16,5%
2024	1º Sem	Logística	Noturno	190	248	30,5%	248	30,5%	186	-2,1%
2024	1º Sem	Polímeros	Noturno	68	54	-20,6%	36	-47,1%	98	44,1%
2025	1º Sem	Desenvolvimento de Software	Noturno	282	431	52,8%	180	-36,2%	294	4,3%
2025	1º Sem	Fabricação Mecânica	Noturno	97	101	4,1%	106	9,3%	55	-43,3%
2025	1º Sem	Informática para Negócios	Noturno	105	118	12,4%	121	15,2%	78	-25,7%
2025	1º Sem	Logística	Noturno	120	241	100,8%	240	100,0%	223	85,8%
2025	1º Sem	Polímeros	Noturno	42	19	-54,8%	17	-59,5%	89	111,9%

## Análises Gerais

**Em 2020 e 2021, a pandemia e o ensino remoto impulsionaram inscrições, sobretudo em Informática (manhã), Logística e Gestão EAD.**

**EAD apresentou boa adesão, talvez seja o momento de ampliar o investimento. (diversificação de cursos).**

**Fabricação Mecânica apresenta variações acentuadas, indica possível influência de fatores externos (mercado de trabalho, sazonalidade)**

## Análises dos Modelos

**Holt mostrou tendencia de suavizar os dados, útil para séries lineares, mas comprometeu a precisão em variações acentuadas.**

**Prophet foi consistente em cenários alteração gradual, mas teve dificuldade em variações bruscas.**

**ARIMA em alguns casos se aproximou muito dos valores reais, mas a baixo volume de dados comprometeu o resultado final.**

**Em alguns casos os modelos divergiram fortemente entre si, pode ser usado um modelo de ensemble (combinação de previsões)**

# Conclusão

**Os 3 modelos (Prophet, Holt e ARIMA) apresentaram diferenças significativas, indicando necessidade de ajustes e calibração.**

**Sugere-se revisar os parâmetros e considerar abordagens combinadas para aumentar a precisão das previsões.**