



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA



DIPARTIMENTO  
**MATEMATICA**

# Italian cinema analysis and forecasting

---

Simone De Renzis

Matteo Bergamaschi

Davide Varotto

• • • • • • • • • • December 1, 2023

Business Economics and Financial Data



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA



DIPARTIMENTO  
MATEMATICA

# INTRODUCTION

# Aim of the work

- Analyze Cinema revenues in Italy
  - How Covid-19 pandemic impacted?
  - Was there a trend already?
- Does the streaming industry play a role of competition?
- How Italian and foreign productions impact the revenues?
- What can we say about the future?



- The film chain has always been a great importance industry that creates a business for over than 4 billion with 2.000 active companies.
- These companies, according to some studies, are increasing revenues by 3%/6% every years.

But who dominates this business?

- Only Cinema sector is worth 600 million and employs over 300 companies
- It's confirmed as the weak link in the supply chain, suffering severe competition from TV and the internet.

# Dataset

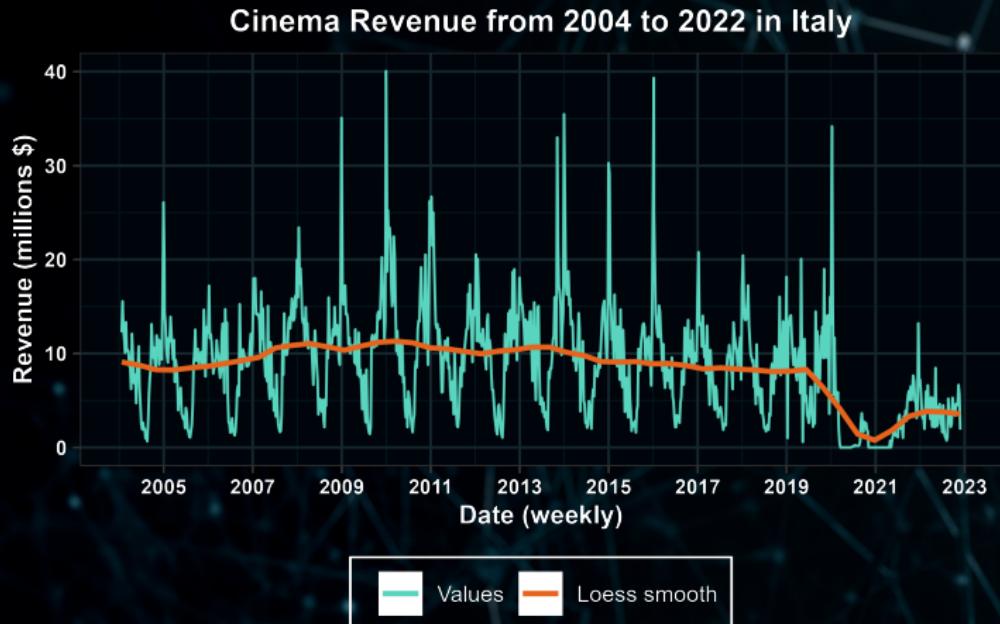
- **Box Office Mojo:** website that tracks box-office revenue, owned by Amazon
- For every country, weekly box office data
- This is source of our data: R script to *webscrape* HTML tables and convert them into data frames

The screenshot shows the Box Office Mojo homepage with a dark theme. The main navigation menu includes Domestic, International, Worldwide, Calendar, All Time, Showdowns, and Indices. A search bar at the top right contains the placeholder "Search for Titles". Below the menu, a section titled "Italian Box Office Weekends For 2022" displays a table of box office data. The table has columns for Dates, Top 10 Gross, %± LW, Overall Gross, %± LW, Releases, #1 Release, and Week. The data shows several weekends in 2022, with notable entries for the World Cup in Qatar.

Dates	Top 10 Gross	%± LW	Overall Gross	%± LW	Releases	#1 Release	Week
Dec 29-Jan 1, 20...	\$10,666,124	+46.1%	\$11,043,412	+48.7%	48	Avatar: The Way of Water	52
Dec 22-25	\$7,302,019	-27.4%	\$7,426,323	-28.3%	26	Avatar: The Way of Water	51
Dec 15-18 World Cup (Qatar)	\$10,064,464	+190.1%	\$10,351,428	+132.7%	25	Avatar: The Way of Water	50
Dec 8-11 World Cup (Qatar)	\$3,468,979	+2.9%	\$4,448,355	+1%	56	Puss in Boots: The Last Wish	49
Dec 1-4 World Cup (Qatar)	\$3,371,140	-6.5%	\$4,405,741	+1.3%	49	Vicini di casa	48

# Time-series

- Weekly revenue in Italy from 2004 to 2022
- Strong seasonality, weak trend
  - Covid-19 had a strong impact





UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA

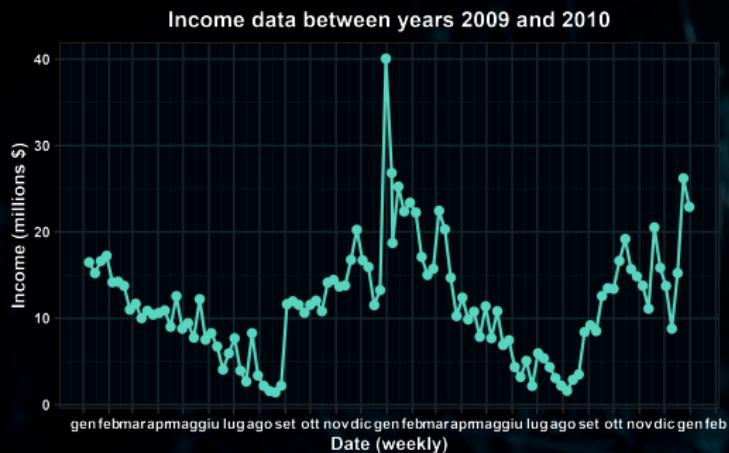


DIPARTIMENTO  
MATEMATICA

# BEFORE COVID

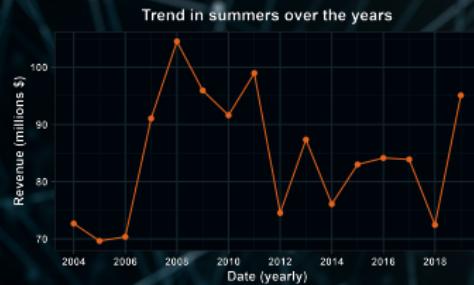
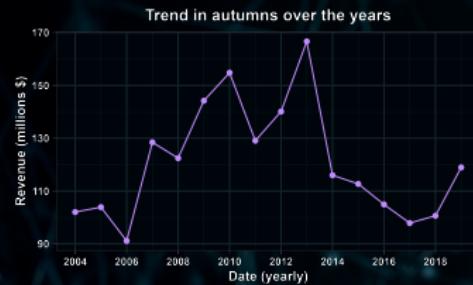
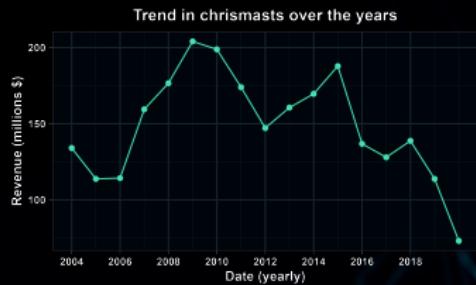
# Most valuable weeks

- Peaks of revenue are shown during Christmas times
  - In the weekend 24-27 December 2009, *Natale a Beverly Hills* amounted for a revenue of over 13 millions alone
- By taking a mean over the years, there are some weeks in which revenues are much higher



# Is there a trend?

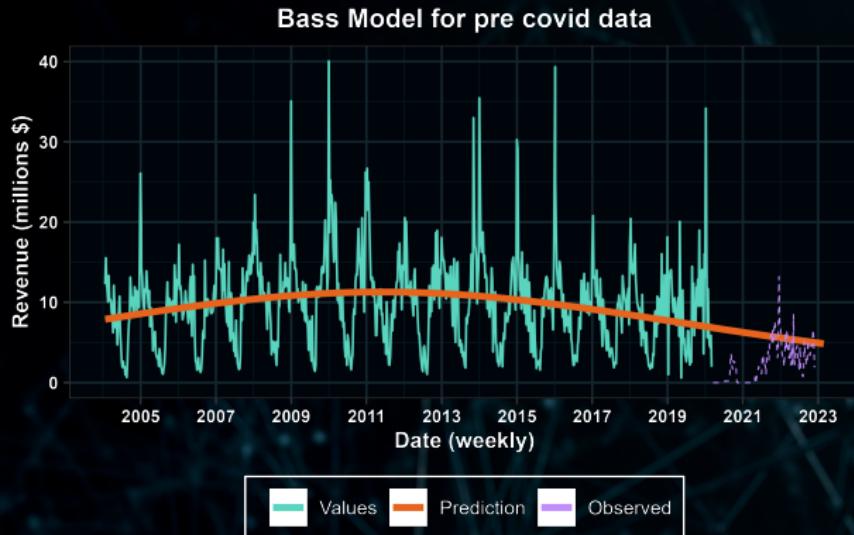
- Let's see if, over the years, different periods of the year present a trend



- Christmas weeks are the most valuable ones, they seem to present a descending trends over the last few years

## Bass model

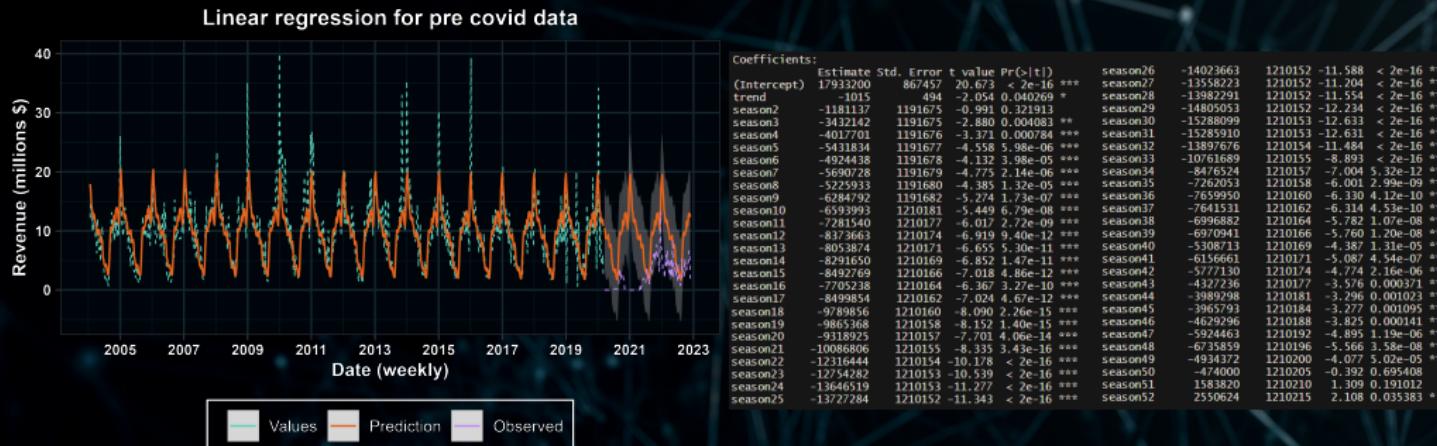
- Bass model seems to capture a peaking and then descending trend, even before the covid-19 pandemic



	Estimate	Std.Error	Lower	Upper	p-value
m	1.088234e+10	4.629262e+07	1.079161e+10	1.097307e+10	0 ***
p	7.244221e-04	1.770733e-06	7.209515e-04	7.278927e-04	0 ***
q	2.489178e-03	2.373916e-05	2.442650e-03	2.535706e-03	0 ***

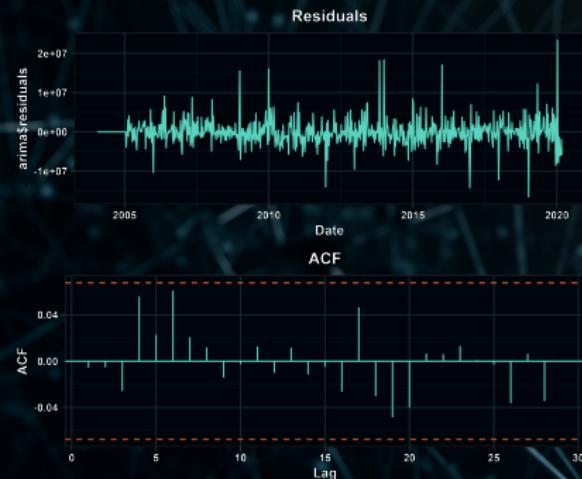
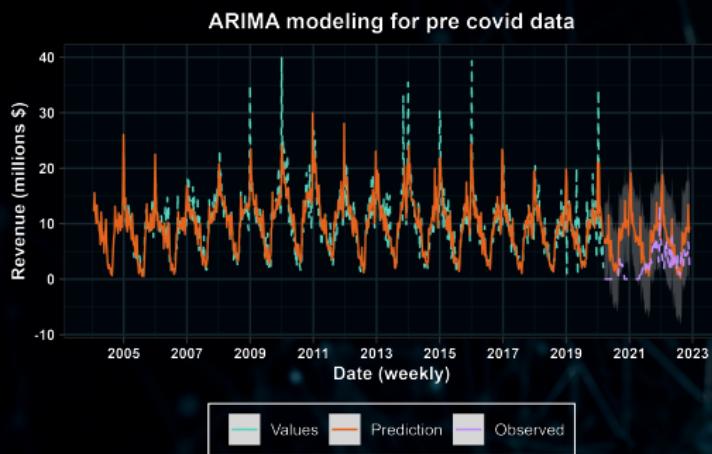
# Prediction - Linear model

- Linear regression model with seasonality and trend
- Seasonality is captured well, the trend not so much



# Prediction - ARIMA

- ARIMA(3, 1, 1)(0, 1, 1)<sub>52</sub> parameters obtained manually by differencing + looking at ACF and PACF
  - Confirmed with Auto Arima
- Capture nicely the trend and seasonality
- Nice residuals and autocorrelation
- Of course, Covid-19 impact is unforeseeable





UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA



DIPARTIMENTO  
MATEMATICA

# COVID-19

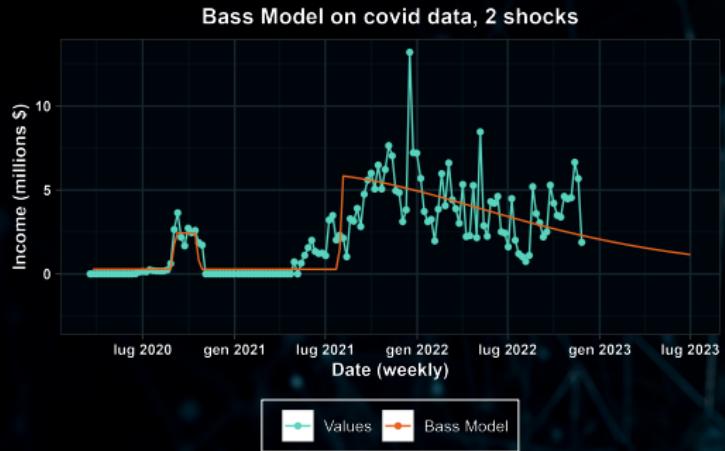
# Covid-19 pandemic

- Global economy has been overwhelmed by this emergency
- In the film industry it further widened the gap that was already building between the streaming services and the cinema industry



# Bass model with shocks

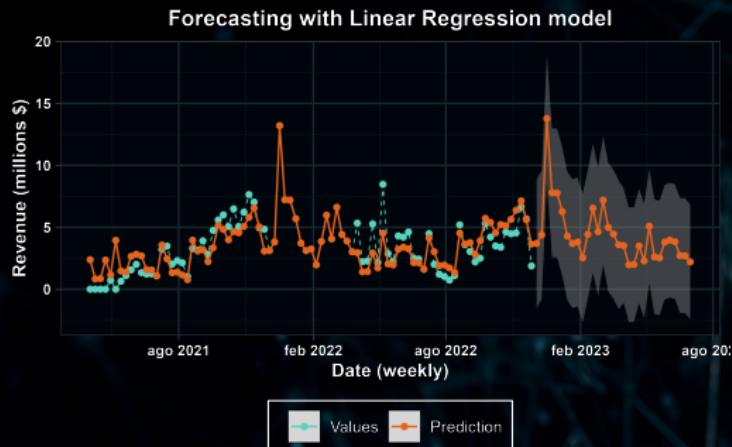
- During lockdowns, cinemas in Italy were closed
- These are shocks that deeply impacted the revenues
  - We model the behaviour with a Bass model
  - Estimates are significant
- Even after reopening, revenues struggle to recover



	Estimate	Std.Error	Lower	Upper	p-value	<chr>
m	4.170097e+08	2.464532e+07	3.687058e+08	4.653136e+08	7.27e-35	***
p	1.428404e-02	5.771440e-04	1.315286e-02	1.541522e-02	1.98e-51	***
q	1.185987e-02	5.015851e-03	2.028984e-03	2.169076e-02	1.95e-02	*
a1	-5.471214e-01	1.989943e-01	-9.371432e-01	-1.570997e-01	6.81e-03	**
b1	7.178663e+01	4.424611e-01	7.091942e+01	7.265384e+01	8.21e-154	***
c1	-9.529712e-01	9.421981e-03	-9.714379e-01	-9.345044e-01	6.06e-127	***
a2	2.431484e+01	1.634445e+00	2.111138e+01	2.751829e+01	5.25e-30	***
b2	3.124996e+01	1.540500e+00	2.823064e+01	3.426929e+01	2.16e-42	***
c2	3.644348e-01	1.382680e-01	9.343442e-02	6.354352e-01	9.40e-03	**

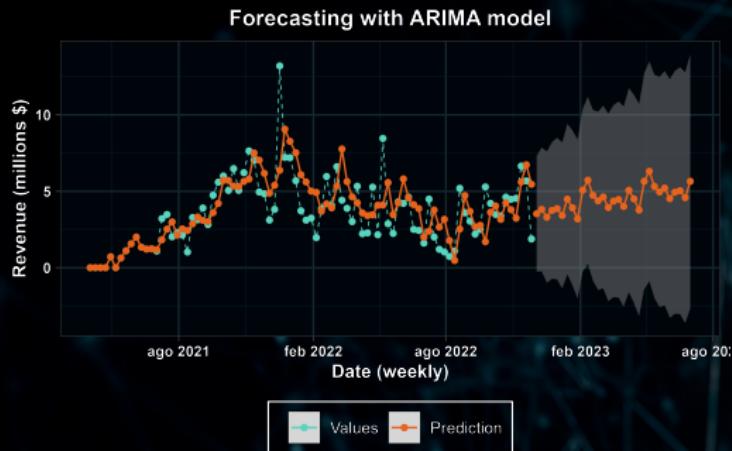
# Forecasting - Linear model

- The linear model doesn't provide meaningful forecasting
- Residuals have a shape



# Forecasting - ARIMA

- ARIMA(2, 1, 1)(0, 1, 1)<sub>52</sub> seems to predict a slightly increasing trend for the future
  - Residuals are nicely shaped and not too much autocorrelated





UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA

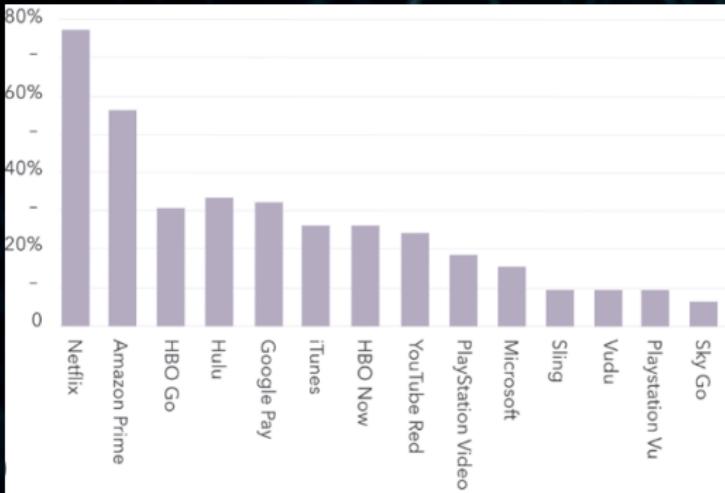


DIPARTIMENTO  
MATEMATICA

# STREAMING

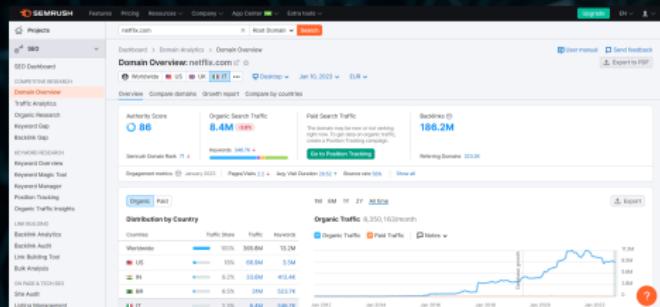
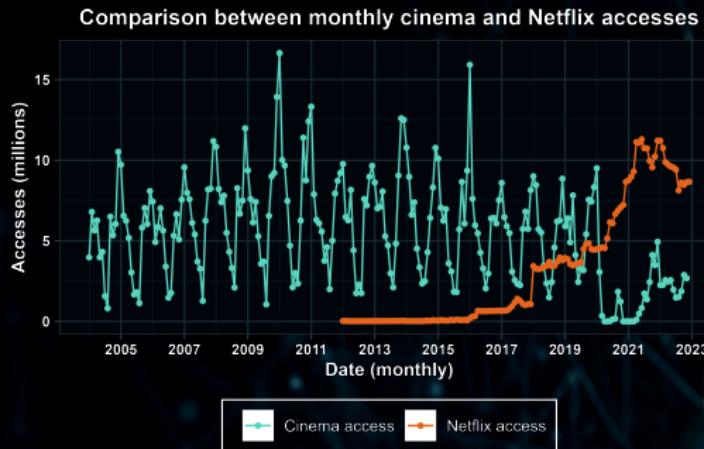
# Streaming Industry

- The film business is dominated by global players
- The giants of on demand, Netflix and Amazon Prime above all, are dominating, changing the rules of the game
- Only Netflix is estimated to have a turnover of 400 million.



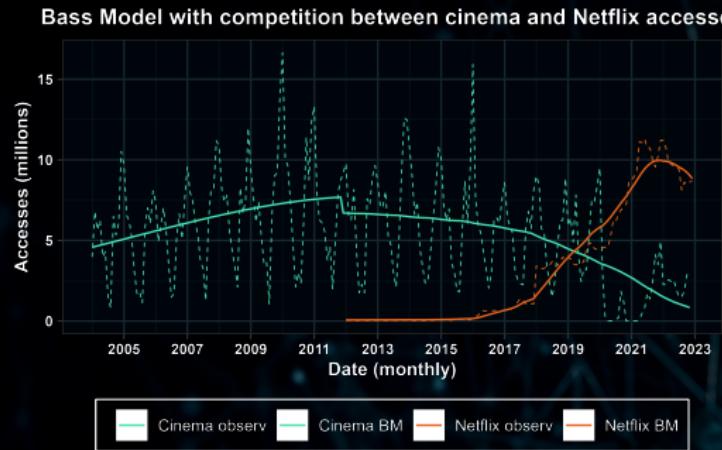
# Netflix competition

- [semrush.com](https://semrush.com) is a platform used for keyword research and online ranking data, including metrics such as search volume
  - We extracted monthly data of access to the *Netflix* streaming service
- To bring into the same scale, we divided the revenues by 7, the cost of an average cinema ticket in Italy



# Netflix competition

- Does Netflix play a role as a competitor?
- The diffusion model for competition seems to capture the behaviour
  - The estimates are significant
- The decline started even before the pandemic, but is especially visible during Covid-19



	Estimate	Std.Error	Lower	Upper	p value <hr>
ma	1.784645e+09	3.410358e+08	1.116227e+09	2.453062e+09	1.04e-06 ***
p1a	2.552203e-03	4.445343e-04	1.680932e-03	3.423475e-03	1.20e-07 ***
q1a	1.170134e-02	1.492935e-03	8.752173e-03	1.462744e-02	7.82e-12 ***
mc	1.323315e+09	6.115622e+07	1.203451e+09	1.443179e+09	7.89e-60 ***
p1c	5.072163e-03	4.377748e-04	4.214140e-03	5.930186e-03	2.97e-25 ***
p2	4.819901e-05	3.949017e-04	-7.257941e-04	8.221922e-04	9.03e-01
q1c	-1.264363e-02	5.660151e-03	-2.374000e-02	-1.547266e-03	2.64e-02 *
q2	1.071673e-01	1.394425e-02	7.983704e-02	1.344975e-01	3.23e-13 ***
delta	1.701632e-02	6.990152e-03	3.313205e-03	3.071943e-02	1.56e-02 **
gamma	1.073682e-01	1.549290e-02	7.700268e-02	1.377337e-01	3.38e-11 ***



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA



DIPARTIMENTO  
MATEMATICA

# PRODUCTION NATIONALITY

# Italian and foreign productions

- Does the nationality of the movies play a role in the revenues?
- For every weekend of every year
  - For every movie airing: get the link of the distribution company
  - From the link, we retrieve the company code, (e.g. co0226183) can be searched in the *IMDb* (Internet Movie Database) website, which shows the nationality
  - Done automatically with an R script

**Box Office Mojo** by IMDbPro

Search for Titles

Domestic International Worldwide Calendar All Time Showdowns Indices

Italian 2022 Weekend 50  
December 15-18, 2022

Weekend Chart ▾

« Week Year Week »

Key: New This Week

Rank	LW	Release	Gross	%t LW	Theaters	Change	Average	Total Gross	Weeks	Distributor
1	-	Avatar: The Way of Water	\$8,991,132	-	595	-	\$15,111	\$10,567,661	1	Walt Disney Studios Motion Pictures
2	1	Puss in Boots: The Last Wish	\$462,407	-66%	490	-28	\$943	\$2,632,649	2	Universal Pictures International (UPI)
3	2	Vicini di casa	\$165,280	-70.5%	332	-177	\$497	\$1,584,374	3	Medusa Distribuzione

IMDb

☰ Menu Aziende ▾ Cerca su IMDb

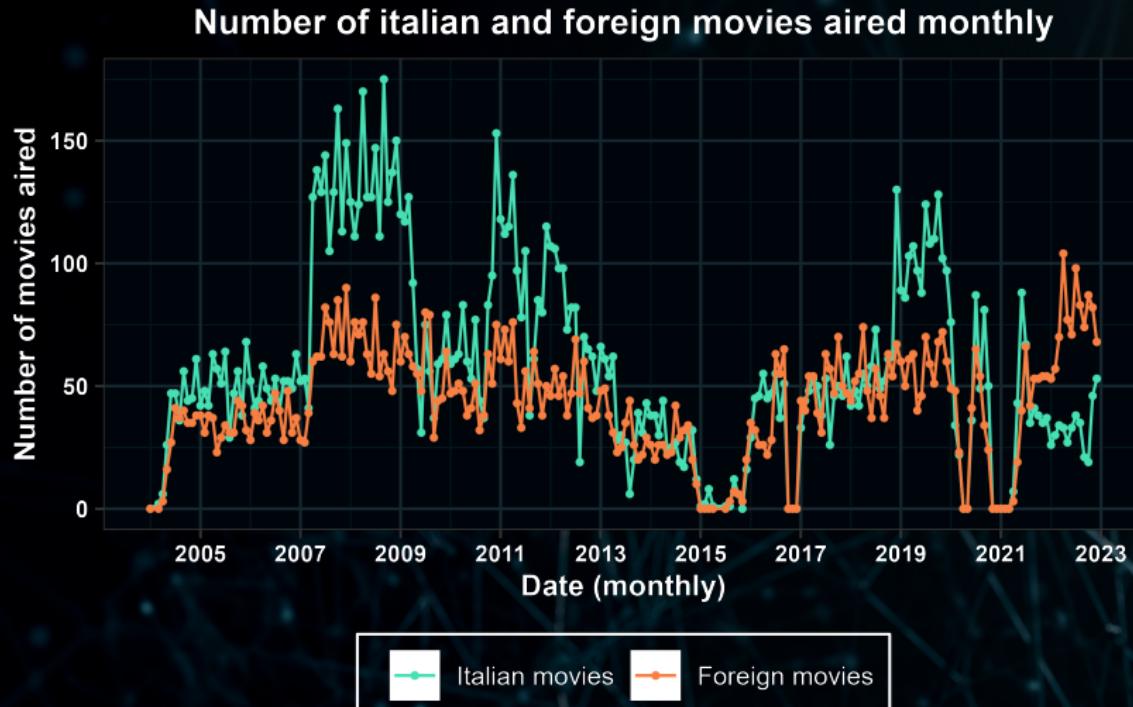
Cerca “co0226183”

Azienda

Walt Disney Studios Motion Pictures  
Stati Uniti · Distributore

# Productions

We obtained this plot





UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA

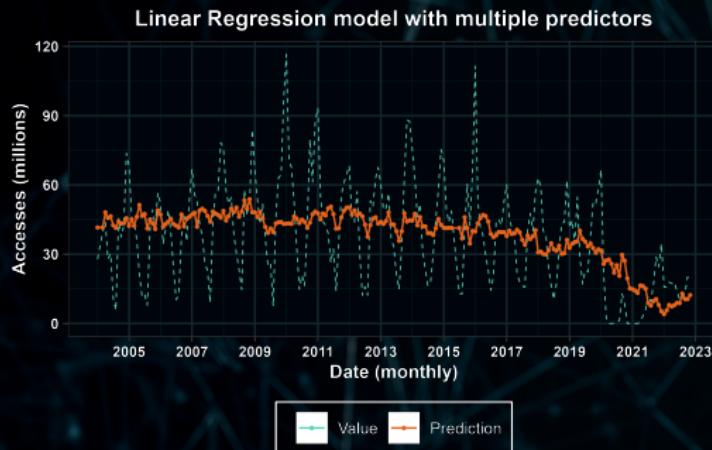


DIPARTIMENTO  
MATEMATICA

# MULTIPLE PREDICTORS

# Linear Regression with multiple predictors

- We created a dataset having two predictors:
  - Number of Netflix accesses
  - Ratio of italian and foreign movies
- Explanatory variable: number of cinema accesses
- Linear regression seems to capture a relationship

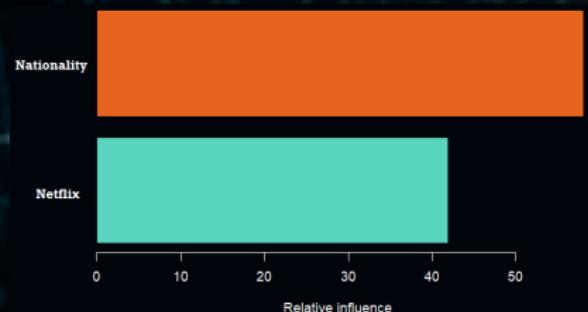
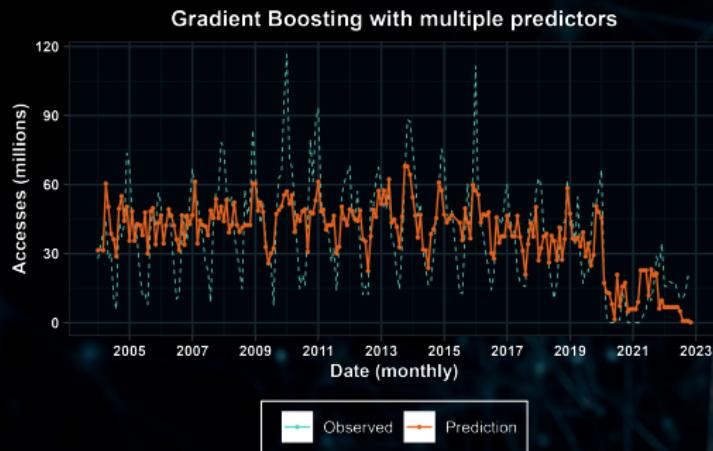


Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	3.493e+07	3.922e+06	8.905	< 2e-16 ***
cinema_data\$perc	6.616e+06	2.493e+06	2.654	0.00853 **
cinema_data\$netflix_access	-3.049e+00	4.264e-01	-7.149	1.26e-11 ***

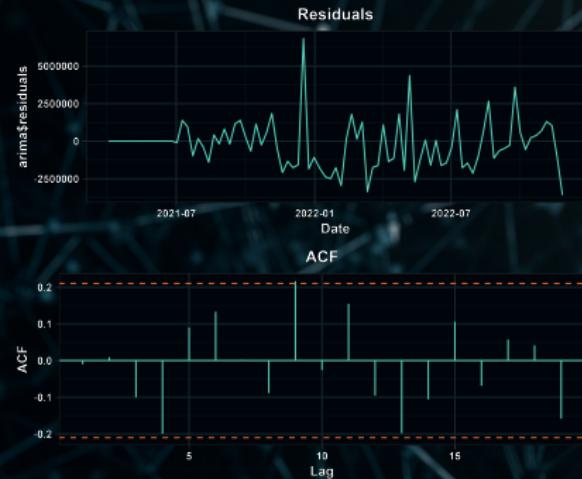
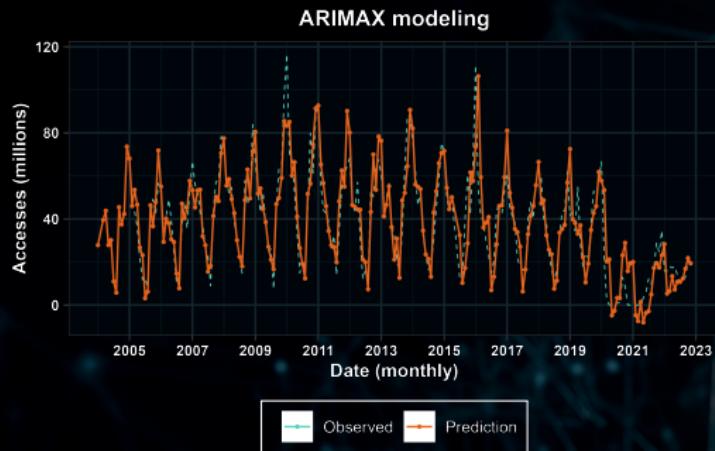
# Gradient Boosting with multiple predictors

- The same dataset is used for the Gradient Boosting model
- The relationship is confirmed, both predictors have a significant role



# ARIMAX

- ARIMAX seems to have a nice fit, confirming the relationship between the predictors and the explanatory variables



- Cinema revenues in Italy already had a slight descending trend before Covid-19
  - The streaming industry is playing a role in this
- The pandemic impacted heavily on cinema access
  - After the pandemic, streaming services are being highly used, while cinema is still struggling to recover
- Accesses at the cinema seem to be higher when there are more Italian movies