

# Progetto di Informatica Umanistica

**1. Titolo:** Analisi dell'industria cinematografica dal 1980 al 2001

**2. Autori:**

- Alessandro Dalla Riva [151881@spes.uniud.it](mailto:151881@spes.uniud.it) 151881;
- Demien Propedo [153260@spes.uniud.it](mailto:153260@spes.uniud.it) 153260;
- Matteo Tavano [154255@spes.uniud.it](mailto:154255@spes.uniud.it) 154255.

**3. Abstract:**

Nel nostro studio sull'industria cinematografica, abbiamo esaminato l'impatto dell'aumento del PIL mondiale sugli incassi e il budget dei film. Abbiamo analizzato i generi maggiormente prodotti e quelli che hanno incassato di più. Infine, ci siamo concentrati su alcune produzioni in cui ha recitato Tom Cruise.

**4. Introduzione:**

Il presente progetto si propone di esplorare la correlazione tra lo sviluppo dell'industria cinematografica e l'aumento del Prodotto Interno Lordo (PIL) mondiale nel periodo compreso tra il 1980 e il 2001. In questo arco di tempo, il cinema ha svolto un ruolo cruciale nell'intrattenimento globale e nella diffusione della cultura popolare. Una delle chiavi del nostro studio sarà l'analisi dei generi cinematografici che hanno registrato i maggiori incassi durante questo periodo, offrendo una prospettiva sulla preferenza del pubblico e sulle tendenze cinematografiche predominanti. Allo stesso tempo, esamineremo anche i generi che hanno visto la produzione del maggior numero di film, rivelando le dinamiche di mercato e l'offerta di contenuti. Infine, daremo uno sguardo ai risultati ottenuti dai film interpretati da Tom Cruise nel medesimo periodo, concentrandoci sui generi più prolifici e quelli che hanno generato i profitti più consistenti, per comprenderne l'impatto sul panorama cinematografico globale.

**5. Fonti letterarie:**

Nel corso del nostro studio, ci siamo basati ampiamente sul libro di Enrico Maria Mastroddi, "Il cinema digitale in Europa". Analisi delle potenzialità del digitale nel

settore cinematografico”, il quale ha svolto un ruolo fondamentale nel plasmare la nostra ricerca. Quest'opera ci ha fornito preziose prospettive sulle potenzialità del cinema digitale nell'industria cinematografica e ha contribuito a definire molti degli aspetti centrali che abbiamo esaminato nel corso del nostro lavoro. Gli insight e le analisi presenti in questo libro ci hanno permesso di approfondire la nostra comprensione dei cambiamenti e delle sfide affrontate dall'industria cinematografica nel contesto digitale, fornendoci una solida base per le nostre ricerche.

Inoltre, abbiamo preso spunto dalla ricerca condotta da Choujun Zhan, Jianjin Li e Wei Jiang, “An Empirical Investigation on Movie Industry from 1980 to 2018”.

Tuttavia, abbiamo concentrato la nostra indagine su un periodo specifico, che va dagli anni '80 al 2001, e ci siamo focalizzati anche su temi e aspetti differenti rispetto a quelli affrontati nell'articolo di riferimento. Questa scelta ci ha permesso di esplorare in dettaglio una fase chiave dell'industria cinematografica, prendendo in considerazione fattori e dinamiche specifiche di quell'epoca.

## **6. Metodologia e dati:**

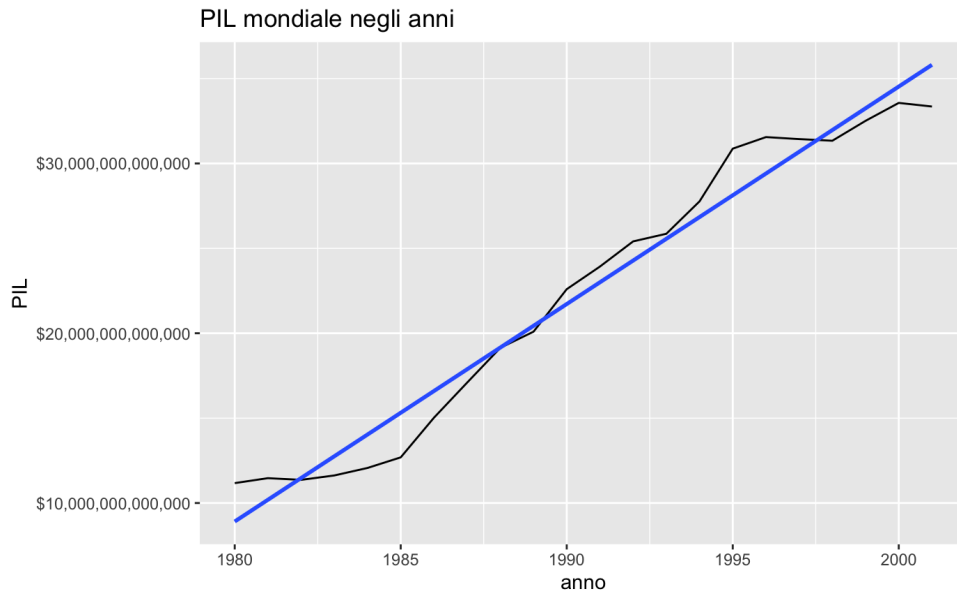
Per condurre un'analisi sulla crescita del settore cinematografico nel corso degli anni utilizzando RStudio, abbiamo adottato un approccio metodologico basato sulla raccolta e l'elaborazione dei dati. In primo luogo, abbiamo acquisito due dataset pertinenti alle informazioni relative ai film (nome, rating, genere, anno..) e relative al gdp (prodotto interno lordo) delle nazioni produttrici. Questi dati sono stati importati in RStudio e preparati per l'analisi, compreso il trattamento di valori mancanti e la pulizia dei dati.

Successivamente, abbiamo utilizzato diverse librerie di R, come *ggplot2*, *readr*, *dplyr*, e *scales* per creare visualizzazioni grafiche accattivanti dei dati. Questi grafici ci hanno aiutato a identificare tendenze di crescita nel numero di film prodotti, nell'incasso al botteghino, nei generi cinematografici più popolari e altro ancora. Abbiamo anche applicato tecniche di analisi temporale per esplorare le fluttuazioni nel settore nel corso degli anni.

RStudio ci ha fornito un ambiente interattivo e potente per eseguire analisi statistiche avanzate, calcolare indicatori di crescita, e generare visualizzazioni chiare e informative. Questa metodologia ci ha permesso di tracciare la crescita del settore cinematografico nel tempo, identificare punti di svolta significativi e sviluppare una

comprensione più approfondita delle dinamiche che guidano l'industria cinematografica.

## 7. Visualizzazione #1:



La visualizzazione mostra l'aumento del PIL mondiale nel tempo. Ciò significa che, dal 1980 al 2001, c'è stato uno sviluppo economico che ha portato ad una maggiore disponibilità economica.

Abbiamo utilizzato un grafico a linee con una linea di tendenza. Questo tipo di grafico è ideale per rappresentare dati che devono essere confrontati nel tempo, in questo modo la lettura è facilitata.

Sull'asse delle ascisse troviamo la variabile "anno", mentre su quello delle ordinate troviamo la variabile "PIL".

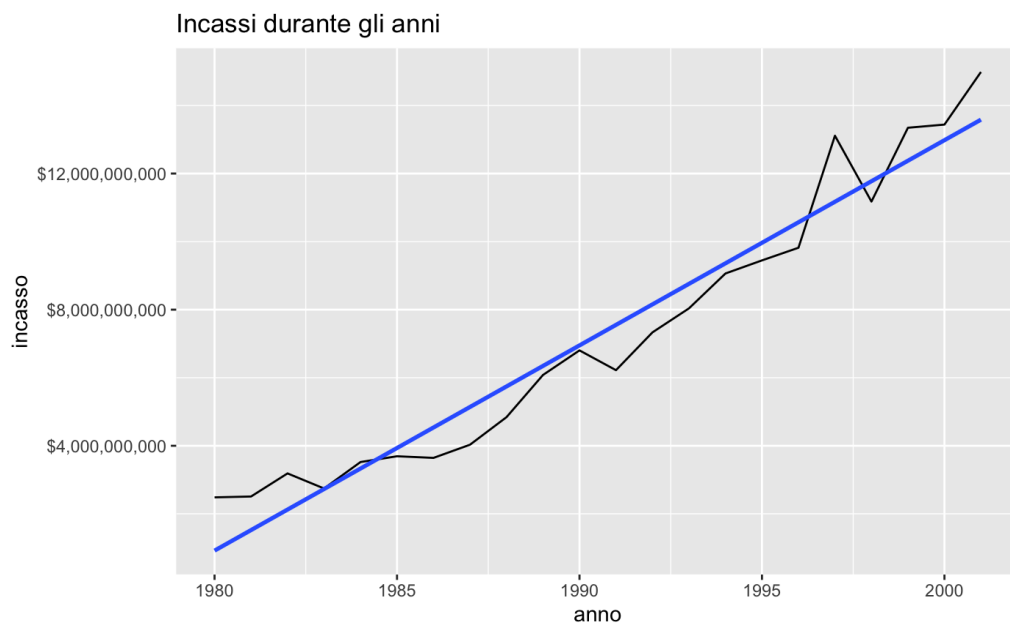
Questo grafico mostra che in 10 anni, dal 1985 al 1995, c'è stata una crescita vertiginosa del PIL mondiale.

Tramite il grafico possiamo vedere e confrontare in modo più chiaro come è aumentato il PIL nel tempo.

Il grafico evidenzia chiaramente l'andamento generale del PIL mondiale nel periodo compreso tra il 1980 e il 2001. Mostra una linea di tendenza (in blu) che rappresenta la direzione generale dei cambiamenti nel PIL mondiale nel tempo. Tuttavia, il grafico non fornisce dettagli sulle cause di questa crescita e potrebbe non essere adatto a rappresentare periodi di tempo molto ampi. La scala dell'asse delle ordinate potrebbe essere fuorviante se non viene interpretata attentamente. Queste problematiche

derivano in parte dalle limitazioni intrinseche al tipo di visualizzazione scelto, che può non essere in grado di rappresentare dettagli complessi o fornire spiegazioni approfondite. Pertanto, per una comprensione completa delle dinamiche di crescita del PIL mondiale e delle sue cause, potrebbe essere necessaria un'analisi più approfondita che vada oltre il grafico e consideri i fattori economici, politici e sociali che hanno contribuito a questa crescita.

## 8. Visualizzazione #2:



La visualizzazione mostra l'aumento degli incassi dei film nel tempo. Possiamo notare una chiara correlazione con l'aumento del PIL. Ciò è dovuto dal fatto che le persone, disponendo di un reddito maggiore, avevano una disponibilità economica maggiore per le spese extra, nel nostro caso l'intrattenimento.

Abbiamo usato sempre un grafico a linee con una linea di tendenza.

Le variabili che sono rappresentate sono le seguenti: sull'asse delle ascisse troviamo l'anno e sull'asse delle ordinate troviamo l'incasso dei film in quell'anno.

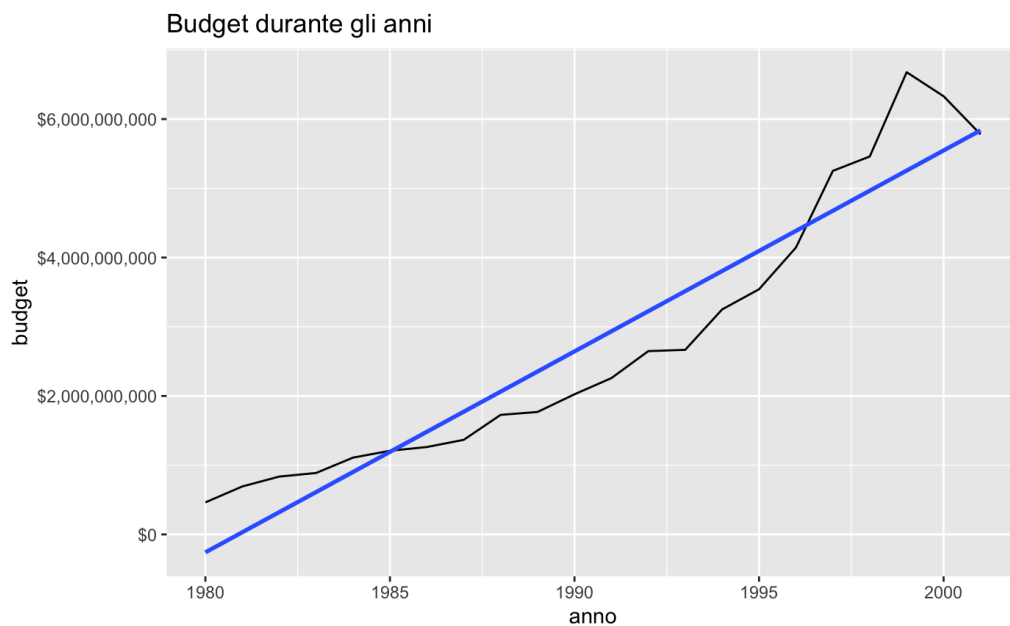
Questo grafico ci mostra una netta crescita degli incassi nel periodo che va dal 1986 al 1997.

Il grafico rende immediatamente evidenti le fluttuazioni negli incassi annuali e fornisce una visione chiara delle tendenze nel corso degli anni. Questi dettagli potrebbero non emergere con altrettanta chiarezza se si osservassero direttamente i dati tabellari.

Tuttavia, il grafico non fornisce dettagli sui singoli film o sugli eventi che potrebbero aver influenzato gli incassi in determinati anni.

Le problematiche potrebbero essere dovute al numero di anni nel dataset e alla dimensione del grafico, ma sono anche legate alle limitazioni intrinseche al tipo di visualizzazione scelto. Per ottenere dettagli sui film specifici o sugli eventi, sarebbe necessario esaminare i dati di base o creare altre visualizzazioni specifiche.

## 9. Visualizzazione #3:



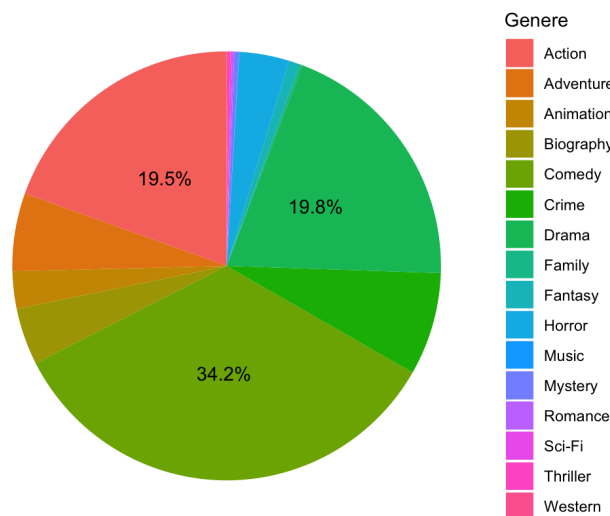
La visualizzazione mostra che il budget è aumentato di conseguenza al crescere dell'industria cinematografica, con un picco associato all'arrivo del nuovo millennio e una lieve decrescita nei primi anni 2000.

Ricordiamo che gli anni '90 dell'industria cinematografica hanno prodotto dei capolavori come "Titanic", "Star Wars" e "Jurassic Park". Il budget associato a tali spettacoli era piuttosto elevato, per favorire effetti speciali e scene rivoluzionarie. Abbiamo scelto di utilizzare un grafico a linee per rappresentare l'andamento del budget dei film nel corso degli anni. Questa scelta è stata motivata dalla necessità di visualizzare in modo efficace come una variabile continua, ovvero il budget totale, cambiasse nel tempo, ossia negli anni. Il grafico a linee è stato preferito in quanto permette di evidenziare le variazioni nel budget nel corso degli anni in modo continuo, rendendo più agevole l'osservazione di eventuali trend temporali. Per fornire ulteriori informazioni sulla tendenza dei dati nel tempo, abbiamo aggiunto una linea di tendenza che, pur non essendo sempre in grado di catturare tutte le

fluttuazioni, aiuta a individuare andamenti generali. Nel grafico, le variabili presentate includono l'asse delle ascisse (anno) e l'asse delle ordinate (budget), espresso in dollari. Questo ci permette di visualizzare chiaramente come il budget totale dei film sia cambiato nel corso degli anni, consentendoci di individuare eventuali aumenti o diminuzioni nel finanziamento e di identificare tendenze temporali significative.

#### 10. Visualizzazione #4:

Numero di film per genere



Nella visualizzazione qui sopra andiamo ad osservare qual è la distribuzione dei film per genere. Possiamo notare che il genere dei film maggiormente prodotti è: “Commedia” 34.2%; “Drammatico” 19.8%; “Azione” 19.5%.

Abbiamo usato un grafico a torta per rappresentare in modo chiaro le proporzioni.

Tramite questo grafico è facilmente visibile quale categoria è più grande rispetto alle altre.

L'unica variabile che troviamo nel grafico è il genere del film.

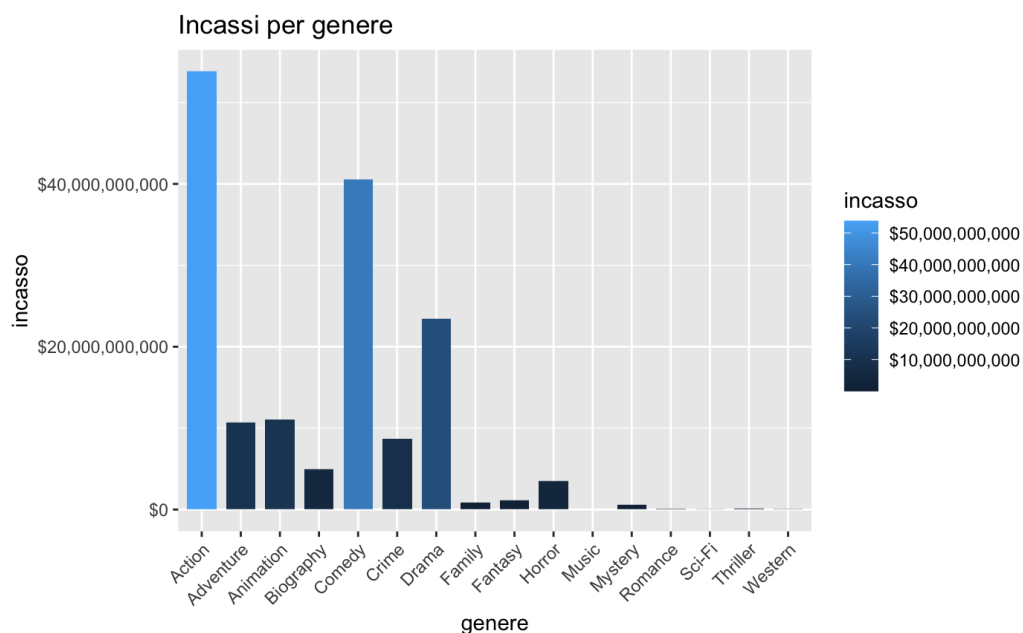
Il grafico permette di visualizzare le percentuali e le proporzioni, che dal dataset non erano chiare. Infatti le abbiamo dovute calcolare tramite delle funzioni.

Tuttavia, ci sono alcune problematiche nel grafico. La sovrapposizione delle etichette percentuali può verificarsi quando ci sono generi con piccole percentuali, rendendo difficile la lettura delle percentuali sui generi meno prodotti. Le etichette percentuali sono posizionate nel centro delle fette, il che può portare a sovrapposizioni o a

difficoltà nella lettura. Per sopperire a questo problema abbiamo inserito una condizione nel codice che ci permette di visualizzare solo le percentuali più alte. Inoltre, il grafico non fornisce dettagli sui nomi dei generi o sul numero esatto di film per ciascun genere.

Queste problematiche sono in gran parte dovute al tipo di grafico a torta utilizzato, che si concentra sulla rappresentazione proporzionale dei generi ma può avere difficoltà a gestire la visualizzazione delle etichette percentuali in caso di fette molto piccole.

## 11. Visualizzazione #5:



Tuttavia, contrariamente al risultato precedente, possiamo notare come il genere “Action” abbia incassato molto di più rispetto agli altri generi, con oltre \$50 miliardi di dollari.

Sebbene il genere “Comedy” sia stato quello che conta più film prodotti, il suo incasso totale si classifica al secondo posto della graduatoria (40\$ miliardi di dollari).

Al terzo posto, troviamo il genere “Drama” che ha incassato oltre \$20 miliardi.

Altri generi hanno incassato cifre relativamente basse rispetto alle categorie precedenti.

La visualizzazione mostra quanto ha incassato ogni genere. Si può facilmente vedere come i generi che hanno incassato di più sono quelli di “Azione”, “Commedia” e

“Drammatici”. Essi corrispondono anche a quelli che sono stati maggiormente prodotti.

Quindi è visibile la correlazione tra la produzione e la richiesta degli spettatori, che porta al conseguente incasso.

Tuttavia, contrariamente al risultato precedente, possiamo notare come il genere “Azione” abbia incassato molto di più rispetto agli altri generi, con oltre \$50 miliardi di dollari.

Sebbene il genere “Commedia” sia stato quello che conta più film prodotti, il suo incasso totale si classifica al secondo posto della graduatoria (40\$ miliardi di dollari). Al terzo posto, troviamo il genere “Drammatico” che ha incassato oltre \$20 miliardi. Altri generi hanno incassato cifre relativamente basse rispetto alle categorie precedenti.

Abbiamo usato un grafico a barre in modo da facilitare la comparazione tra le due variabili, che sono: genere e incasso.

Gli elementi resi più chiari dal grafico includono l'identificazione dei generi cinematografici con gli incassi più elevati. Il grafico fornisce una visione immediata delle differenze di incasso tra i generi.

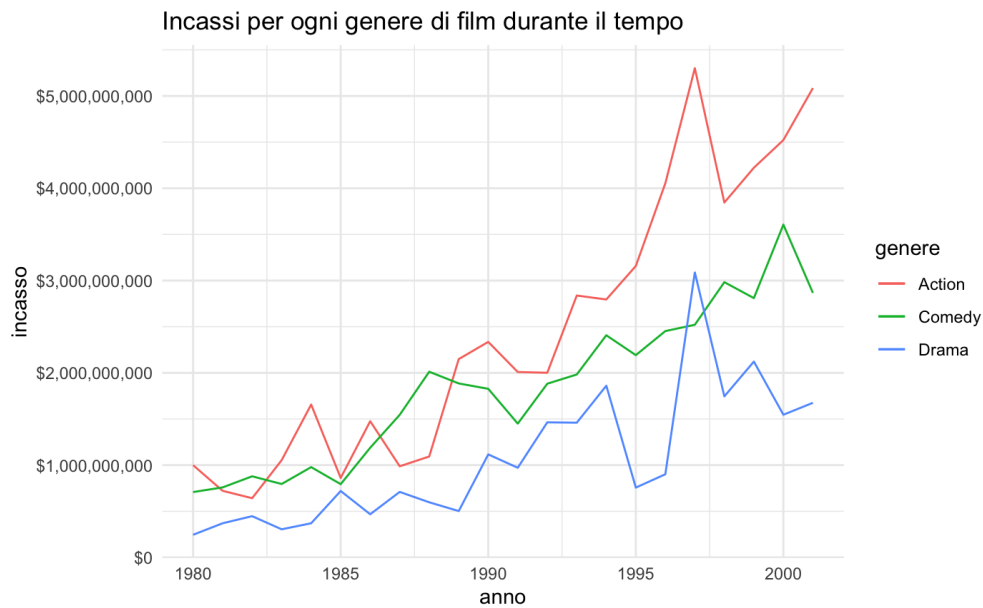
Allo stesso tempo, ci sono alcune problematiche di rappresentazione e interpretazione nel grafico. La sovrapposizione dei nomi dei generi sull'asse x può rendere difficile la lettura, specialmente quando ci sono molti generi nel dataset. Infatti, per risolvere questo problema, abbiamo posizionato le etichette con un angolo di 45°.

Inoltre, il grafico non fornisce dettagli sul numero di film per ciascun genere o sulle ragioni specifiche degli incassi più elevati per alcuni generi.

Le problematiche riscontrate sono in gran parte dovute alle caratteristiche del tipo di grafico utilizzato, che si concentra sulla rappresentazione aggregata dei dati e può presentare difficoltà quando si gestiscono molti generi cinematografici.



## 12. Visualizzazione #6:



La visualizzazione mostra l'andamento degli incassi dei film nei generi "Azione", "Commedia" e "Drammatici" nel corso del tempo (dagli anni '80 al 2001). Ci siamo incentrati solo su questi tre generi avendo visto nella visualizzazione precedente essere quelli con gli incassi maggiori.

Il tipo di grafico utilizzato è un grafico a linee che rappresenta l'andamento temporale degli incassi per ciascun genere.

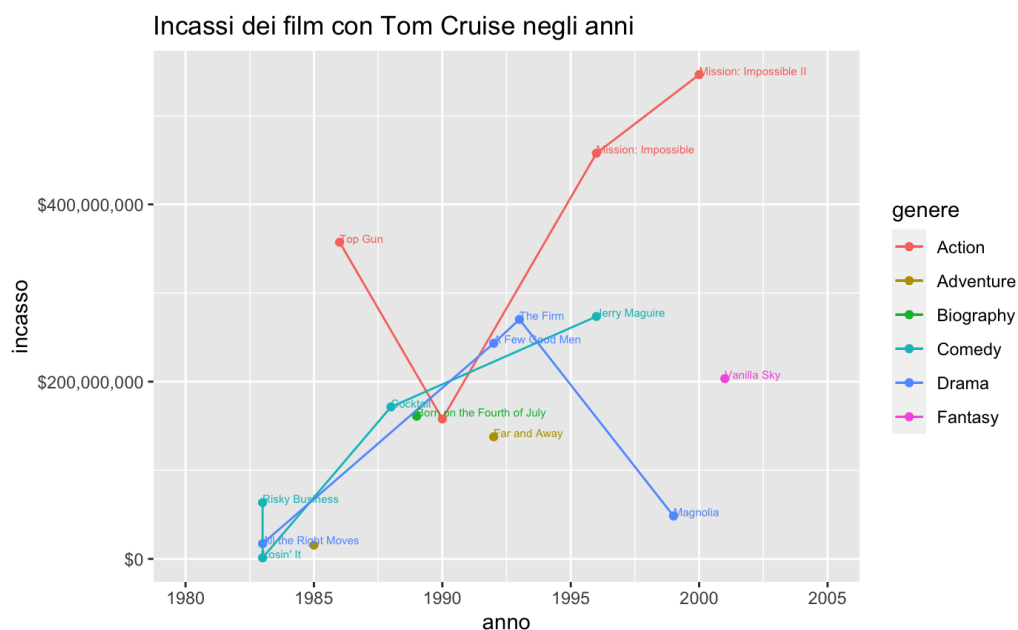
Abbiamo scelto questo tipo di grafico perché è utile per mostrare come gli incassi cambiano nel tempo per ciascun genere, consentendo una comparazione diretta tra i generi.

Le variabili presentate nel grafico sono le seguenti: sull'asse x troviamo il tempo in anni; sull'asse y troviamo l'incasso totale in dollari; come variabile "color" troviamo il genere.

Il grafico mostra come gli incassi dei generi "Azione", "Commedia" e "Drammatici" sono variati nel corso degli anni. Possiamo vedere se ci sono tendenze o fluttuazioni significative negli incassi di questi generi nel tempo.

Nel grafico, potrebbe essere difficile distinguere le linee dei generi se ci sono molte sovrapposizioni o se ci fossero molti generi diversi rappresentati contemporaneamente. Per ovviare a questo problema abbiamo preso in considerazione solo i tre generi di nostro interesse.

### 13. Visualizzazione #7:



In questa visualizzazione possiamo vedere come Tom Cruise, uno degli attori più pagati in quegli anni, abbia riscosso talmente tanto successo, a tal punto da avere gli incassi che possiamo vedere nel grafico, anche grazie al fatto che nella sua carriera ha svolto prevalentemente film di “Azione”, “Commedia” e “Drammatici”. Inoltre, si può notare come i film di “Azione” sono quelli che hanno raggiunto gli incassi più alti, quindi troviamo una correlazione con la visualizzazione n° 5.

Abbiamo scelto questo tipo di grafico perché consente di visualizzare l'andamento temporale degli incassi per diversi generi di film interpretati da Tom Cruise, consentendo una comparazione diretta tra i generi e individuando eventuali tendenze nel tempo.

Le variabili presentate nel grafico sono:

- Asse x: Anno (year)
- Asse y: Incasso totale (gross) in dollari
- Colore: Genere del film (Action, Comedy, Drama)
- Etichette dei punti: Nome del film (name)

Questo grafico ci mostra come gli incassi dei film interpretati da Tom Cruise sono variati nel corso degli anni nei generi “Azione,” “Commedia,” e “Drammatico.” Si può vedere come gli incassi di ciascun genere sono cresciuti o diminuiti nel tempo e se ci sono state tendenze specifiche per ogni genere.

Tuttavia, con l'aggiunta delle etichette dei punti che mostrano il nome del film, la visualizzazione potrebbe diventare caotica, essendoci alcune sovrapposizioni. Gli incassi dei film di Tom Cruise sono visualizzati nel corso degli anni per i generi "Azione," "Commedia," e "Drammatici." Il grafico è abbastanza chiaro nel mostrare come gli incassi siano variati nel tempo per ciascun genere. Tuttavia, ci sono due questioni di rappresentazione e interpretazione da considerare. La prima è la sovrapposizione delle etichette dei film. Poiché sono presenti etichette dei film nei punti corrispondenti agli incassi, si sono verificate delle sovrapposizioni. Questo potrebbe rendere difficile la lettura delle etichette. La seconda è la rappresentazione dei generi. L'utilizzo di linee per rappresentare l'andamento dei generi ha portato anch'esso a sovrapposizioni che potrebbero rendere difficile l'interpretazione del grafico.

#### **14. Conclusioni:**

Dalla nostra analisi dei dati, delle fonti scritte, della metodologia utilizzata e delle visualizzazioni effettuate, emergono diverse conclusioni significative. Innanzitutto, è evidente una forte correlazione tra l'aumento del Prodotto Interno Lordo (PIL) mondiale e l'industria cinematografica nel periodo 1980-2001. L'incremento del PIL ha influenzato positivamente gli incassi e i budget dei film, dimostrando l'interconnessione tra sviluppo economico globale e successo dell'industria cinematografica.

I generi cinematografici più prodotti durante questo periodo sono risultati essere "Azione", "Commedia" e "Drammatici", e questi stessi generi hanno ottenuto gli incassi più elevati. Ciò suggerisce una forte preferenza del pubblico per tali generi e una loro redditività per l'industria.

Un aspetto degno di nota è il ruolo significativo degli attori di successo, come Tom Cruise, che ha dominato il panorama cinematografico con film principalmente di "Azione", "Commedia" e "Drammatici". Questo sottolinea che la sua versatilità nei diversi generi è stata determinante per il suo successo.

Tuttavia, l'analisi dei dati potrebbe non aver esplorato completamente alcuni aspetti, come l'evoluzione delle tecnologie cinematografiche o l'impatto della globalizzazione sull'industria. Pertanto, bisognerebbe eseguire ulteriori ricerche per approfondire questi argomenti.

Nel complesso, i risultati dell'analisi sono chiari riguardo alla correlazione tra PIL e industria cinematografica, ma indicano anche la necessità di ulteriori approfondimenti per comprendere appieno le complesse dinamiche dell'industria cinematografica durante quel periodo.

## 15. Bibliografia:

- *Dataset GDP:*
  - Nome del dataset: World GDP EDA
  - Disponibile su Kaggle:  
<https://www.kaggle.com/code/harisonmwangi/world-gdp-eda/input>
- *Dataset Movies:*
  - Nome del dataset: Movies Updated Data
  - Disponibile su Kaggle:  
<https://www.kaggle.com/datasets/ashishkumarjayswal/movies-updated-data>
- *Libro di Enrico Maria Mastroddi:*
  - Mastroddi, Enrico Maria. "Il cinema digitale in Europa. Analisi delle potenzialità del digitale nel settore cinematografico.", 2008.
  - Pagine 29-39.
- *Articolo di Choujun Zhan, Jianjin Li, Wei Jiang:*
  - Zhan, Choujun, Jianjin Li, e Wei Jiang. "An Empirical Investigation on Movie Industry from 1980 to 2018.", 2018.