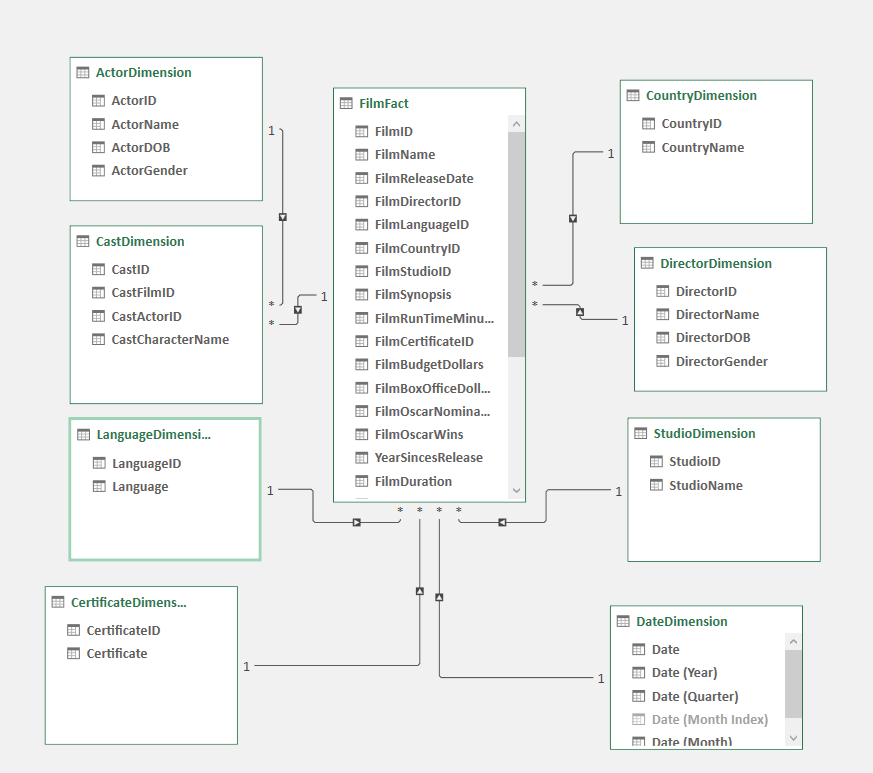
**Data Management II Solution**

**Name:** Zainab Ismail Fareeha

**Registration Number**: CST/19/COM/00305 **Department**: Computer Science

# Task 1: Data Modelling



(a) List of Tables, Columns, Primary Keys and Foreign Keys

|  |  |  |  |
| --- | --- | --- | --- |
| **Table Name** | **Nbr of rows** | **Fact or Dimension** | **Columns** |
| FilmFact | 160 | Fact | |  | | --- | | *FilmID(PK)* | | *FilmName* | | *FilmReleaseDate* | | *FilmDirectorID(FK)* | | *FilmLanguageID(FK)* | | *FilmCountryID(FK)* | | *FilmStudioID(FK)* | | *FilmSynopsis* | | *FilmRunTimeMinutes* | | *FilmCertificateID(FK)* | | *FilmBudgetDollars* | | *FilmBoxOfficeDollars* | | *FilmOscarNominations* | | *FilmOscarWins* | |
| StudioDimension | 20 | Dimension | |  | | --- | | *StudioID(PK)* | | *StudioName* | |
| DirectorDimension | 81 | Dimension | |  | | --- | | *DirectorID(PK)* | | *DirectorName* | | *DirectorDOB* | | *DirectorGender* | |
| ActorDimension | 333 | Dimension | |  | | --- | | *ActorID(PK)* | | *ActorName* | | *ActorDOB* | | *ActorGender* | |
| CastDimension | 813 | Dimension | |  | | --- | | *CastID(PK)* | | *CastFilmID(FK)* | | *CastActorID(FK)* | | *CastCharacterName* | |
| CountryDimension | 4 | Dimension | |  | | --- | | *CountryID(PK)* | | *CountryName* | |
| CertificateDimension | 6 | Dimension | |  | | --- | | *CertificateID(PK)* | | *Certificate* | |
| LanguageDimension | 2 | Dimension | |  | | --- | | *LanguageID(PK)* | | *Language* | |
| Date | 16899 | Dimension | |  | | --- | | *Date(PK)* | |

(b) List of Relationships

|  |  |  |  |
| --- | --- | --- | --- |
| **Table Name** | **Primary Key** | **Table** | **Foreign Key** |
| DirectorDimension | DirectorID | FilmFact | FilmDirectorID |
| StudioDimension | StudioID | FilmFact | FilmFacttudioID |
| CertificateDimension | CertificateID | FilmFact | FilmCertificateID |
| CountryDimension | CountryID | FilmFact | FilmCountryID |
| LanguageDimension | LanguageID | FilmFact | FilmLanguageID |
| CastDimension | CastID | FilmFact | CastFilmID |
| ActorDimension | ActorID | CastDimension | CastActorID |
| DateDimension | Date | FilmFact | FilmReleasedDate |

(c) **Excel-Based Solution:**

**Strengths:**

1. **User-Friendly:** Accessible even to those with limited technical expertise.
2. **Cost-Efficient:** Bundled with Microsoft Office, eliminating additional licensing fees.
3. **Rapid Implementation:** Requires no extra infrastructure, making it ideal for handling small to medium-sized datasets.
4. **Adaptability:** Allows easy adjustments to queries, formulas, and visualizations.

**Weaknesses:**

1. **Scalability Constraints:** Struggles to manage datasets exceeding one million rows.
2. **Performance Limitations:** Consumes significant memory, leading to slowdowns with large datasets.
3. **Manual Intervention:** Relies on manual data updates and refreshes.
4. **Security Limitations:** Lacks comprehensive security features found in enterprise-level solutions.

**SSIS-Based Solution:**

**Strengths:**

1. **Enhanced Performance:** Efficiently processes large datasets with advanced ETL capabilities.
2. **Process Automation:** Supports scheduled workflows, minimizing the need for manual intervention.
3. **Powerful Data Handling:** Provides robust tools for data transformation and cleansing.

Task 2: Data Analysis

## (a) **List of DAX Formulae for Calculated Columns (CC) and Measures (M)**

|  |  |  |  |
| --- | --- | --- | --- |
| **SN** | **Name of CC or M** | **CC or M** | **Formula** |
| 1. | Years since release | CC | =YEAR(TODAY()) - YEAR(FilmFact[FilmReleaseDate]) |
| 2. | Number of PG Cert FilmFact | M | PGCertFilms := CALCULATE(COUNT(FilmFact[FilmID]), CertificateDimension[Certificate] = "PG") |
| 3. | Film duration | CC | = IF(FilmFact[FilmRunTimeMinutes] > 150, "Long", IF(FilmFact[FilmRunTimeMinutes] > 100, "Medium", "Short")) |
| 4. | Profit | M | Profit:=sum([FilmBoxOfficeDollars])-sum(FilmFact[FilmBudgetDollars]) |
| 5. | Director Ranking | M | DirectorRanking:=RANKX(ALL(DirectorDimension), FilmFact[Profit], , DESC) |
| 6. | YTD Profit | M | YTDProfit:=CALCULATE([Profit], DATESYTD(FilmFact[FilmReleaseDate])) |

1. **Answers to Questions**:

|  |  |  |
| --- | --- | --- |
| **Question** | **Answer** | **Evidence** |
| **The actor with 10 casts** | **Samuel L. Jackson** |  |
| **What is the percentage of female casts** | **20.42%** |  |
| **What is the ranking, of director Steven Spielberg?** | **2** |  |

1. **Screenshot of PivotTable 1, PivotChart 1 and PivotTable 2.**

|  |  |
| --- | --- |
| **Pivot Table 1** | **Pivot Chart 1** |
| **Pivot Table 2**  **YTD Profit for 2006** |  |