DBS
Dublin Business School
excellence through learning

CA: Data Warehousing and Business Intelligence

Module Title: Data Warehousing and Business Intelligence

Module Code: B8IT104

Module Leader: Kunwar Madan

Student Name: Manoj Kumar Thangaraj & Shane McDonagh
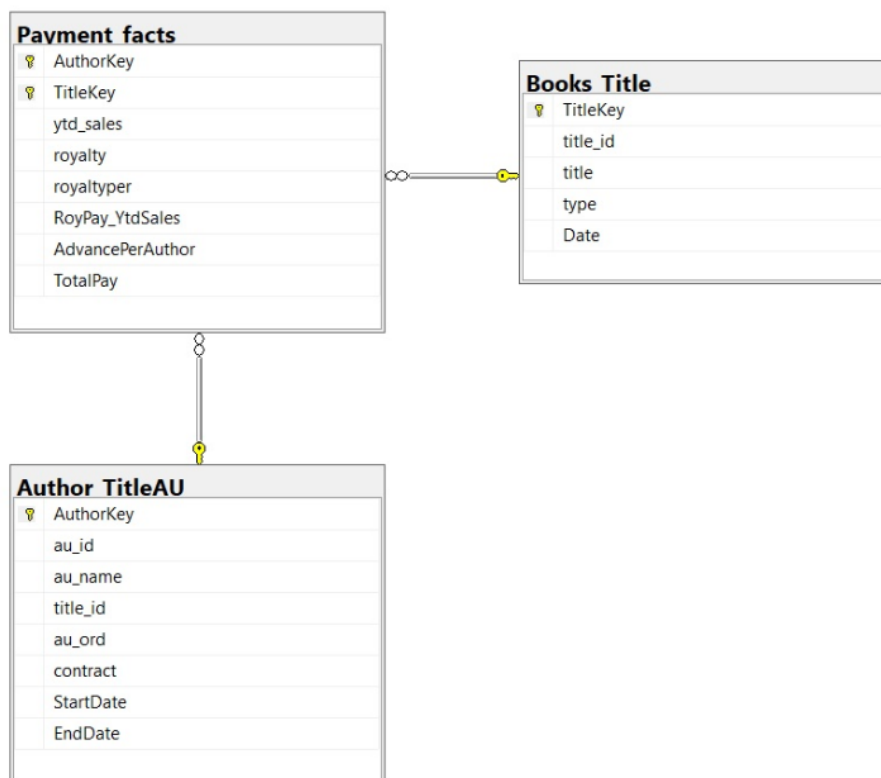
Student Code: 10570753 & 10362717

Contents

1. Project Overview/Scope

We look to create a Data warehouse to produce reports for payments to authors. There are 2 main types of payments: Advance, and Royalties as a percentage of sales. We extracted the information from the publisher's databases via Visual studios, using Kimball star schema. In the relational database, there are 4 tables that include this information. As Authors may write/contribute to more than one book, have identified the atomic grain as individual payments to authors. This allows us to produce reports from this level of detail or in larger forms. These reports may be reproduced over different time periods to identify high earners/ best performing authors/books.

2. Star Schema

Payment facts
- 🔑 AuthorKey
- 🔑 TitleKey
- ytd_sales
- royalty
- royaltyper
- RoyPay_YtdSales
- AdvancePerAuthor
- TotalPay

Books Title
- 🔑 TitleKey
- title_id
- title
- type
- Date

Author TitleAU
- 🔑 AuthorKey
- au_id
- au_name
- title_id
- au_ord
- contract
- StartDate
- EndDate

We extracted the data from 3 tables in the source database: Title, Authors and Titles authors. Roysched is also related to the Authors and payments, but the information is automatically updated in the Titles table, so we did not need it in the schema.

We added a date column to the books dimension table using the SQL function get date, to show when the fields were last updated. In the Author Title AU table, we added contract as a SCD type 2 column to track changes for authors that are in or out of contract. We have for the rest of the data to be updated on each ETL run.

Our facts table has a composite primary key of Title ID and AU ID that have been mapped to the surrogate keys of the dimensional tables, allowing us to track individual payments to authors. This allows us to track individual royalty payments to each author. For the facts table, we created some calculated columns, such as Roypay_ytd as the royalty amount payable and Total pay as royalty plus advance.

3. Reports

We have produced 3 reports for payments, one for advances and one for royalty payments and a combined report for total earnings.

<u>Advances</u>

The first report analyses the advanced paid to authors, also showing the year-to-date sales. We can see Michel DeFrance has received the largest advance.

	Author_Name	book_Name	Sale_to_Date	Advance_Paid
1	Michel DeFrance	The Gourmet Microwave	22246	11250.00
2	Marjorie Green	You Can Combat Computer Stress!	18722	10125.00
3	Cheryl Carson	But Is It User Friendly?	8780	7000.00
4	Sylvia Panteley	Onions, Leeks, and Garlic: Cooking Secrets of the M...	375	7000.00
5	Albert Ringer	Life Without Fear	111	6000.00
6	Livia Karsen	Computer Phobic AND Non-Phobic Individuals: Beha...	375	5250.00
7	Dean Straight	Straight Talk About Computers	4095	5000.00
8	Ann Dull	Secrets of Silicon Valley	4095	4000.00
9	Charlene Locksley	Emotional Security: A New Algorithm	3336	4000.00
10	Reginald Blotchet-Halls	Fifty Years in Buckingham Palace Kitchens	15096	4000.00
11	Sheryl Hunter	Secrets of Silicon Valley	4095	4000.00
12	Anne Ringer	The Gourmet Microwave	22246	3750.00
13	Akiko Yokomoto	Sushi, Anyone?	4095	3200.00
14	Stearns MacFeather	Cooking with Computers: Surreptitious Balance She...	3876	3000.00
15	Abraham Bennet	The Busy Executive's Database Guide	4095	3000.00
16	Burt Gringlesby	Sushi, Anyone?	4095	2400.00
17	Michael O'Leary	Sushi, Anyone?	4095	2400.00
18	Michael O'Leary	Cooking with Computers: Surreptitious Balance She...	3876	2000.00
19	Johnson White	Prolonged Data Deprivation: Four Case Studies	4072	2000.00
20	Marjorie Green	The Busy Executive's Database Guide	4095	2000.00
21	Stearns MacFeather	Computer Phobic AND Non-Phobic Individuals: Beha...	375	1750.00
22	Anne Ringer	Is Anger the Enemy?	2045	1137.50
23	Albert Ringer	Is Anger the Enemy?	2045	1137.50
24	Innes del Castillo	Silicon Valley Gastronomic Treats	2032	0.00
25	Charlene Locksley	Net Etiquette	NULL	NULL

Royalty earned per book and Total earning.

In this report we look at payments to authors based on titles. It shows the royalty and advance payments based on their percentage split and totals them in the total earning column. We can see that Marjorie Green has the highest royalty income.

	Author_Name	book_Name	Sale_to_Date	%_royalty	Amount_of_Royalty_earned	advanceperauthor	Total_earning
1	Marjorie Green	You Can Combat Computer Stress!	18722	100	4493	10125.00	14618
2	Michel DeFrance	The Gourmet Microwave	22246	75	4004	11250.00	15254.25
3	Reginald Blotchet-Halls	Fifty Years in Buckingham Palace Kitchens	15096	100	2113	4000.00	6113
4	Cheryl Carson	But Is It User Friendly?	8780	100	1404	7000.00	8404
5	Anne Ringer	The Gourmet Microwave	22246	25	1334	3750.00	5084.75
6	Dean Straight	Straight Talk About Computers	4095	100	409	5000.00	5409
7	Johnson White	Prolonged Data Deprivation: Four Case Studies	4072	100	407	2000.00	2407
8	Charlene Locksley	Emotional Security: A New Algorithm	3336	100	333	4000.00	4333
9	Abraham Bennet	The Busy Executive's Database Guide	4095	60	245	3000.00	3245.4
10	Innes del Castillo	Silicon Valley Gastronomic Treats	2032	100	243	0.00	243
11	Stearns MacFeather	Cooking with Computers: Surreptitious Balance Sh...	3876	60	232	3000.00	3232.2
12	Ann Dull	Secrets of Silicon Valley	4095	50	204	4000.00	4204.5
13	Sheryl Hunter	Secrets of Silicon Valley	4095	50	204	4000.00	4204.5
14	Marjorie Green	The Busy Executive's Database Guide	4095	40	163	2000.00	2163.6
15	Akiko Yokomoto	Sushi, Anyone?	4095	40	163	3200.00	3363.6
16	Michael O'Leary	Cooking with Computers: Surreptitious Balance Sh...	3876	40	154	2000.00	2154.8
17	Michael O'Leary	Sushi, Anyone?	4095	30	122	2400.00	2522.7
18	Burt Gringlesby	Sushi, Anyone?	4095	30	122	2400.00	2522.7
19	Anne Ringer	Is Anger the Enemy?	2045	50	122	1137.50	1260
20	Albert Ringer	Is Anger the Enemy?	2045	50	122	1137.50	1260
21	Sylvia Panteley	Onions, Leeks, and Garlic: Cooking Secrets of the ...	375	100	37	7000.00	7037
22	Livia Karsen	Computer Phobic AND Non-Phobic Individuals: Beh...	375	75	27	5250.00	5277.75
23	Albert Ringer	Life Without Fear	111	100	11	6000.00	6011
24	Stearns MacFeather	Computer Phobic AND Non-Phobic Individuals: Beh...	375	25	9	1750.00	1759.25
25	Charlene Locksley	Net Etiquette	NULL	100	NULL	NULL	NULL

Total earning by each Author

In this report we calculate the total royalty and advance earned for each Author. We can see that Marjorie Green again has the highest income.

	Au_name	total_advance_earned	royalaty_earned	Total_earning
1	Marjorie Green	12125.00	4656	16781.6
2	Michel DeFrance	11250.00	4004	15254.25
3	Cheryl Carson	7000.00	1404	8404
4	Albert Ringer	7137.50	133	7271
5	Sylvia Panteley	7000.00	37	7037
6	Anne Ringer	4887.50	1456	6344.75
7	Reginald Blotchet-Halls	4000.00	2113	6113
8	Dean Straight	5000.00	409	5409
9	Livia Karsen	5250.00	27	5277.75
10	Stearns MacFeather	4750.00	241	4991.45
11	Michael O'Leary	4400.00	276	4677.5
12	Charlene Locksley	4000.00	333	4333
13	Sheryl Hunter	4000.00	204	4204.5
14	Ann Dull	4000.00	204	4204.5
15	Akiko Yokomoto	3200.00	163	3363.6
16	Abraham Bennet	3000.00	245	3245.4
17	Burt Gringlesby	2400.00	122	2522.7
18	Johnson White	2000.00	407	2407
19	Innes del Castillo	0.00	243	243

4. Conclusion

The analysis on the payments to the authors in the main intend of the project which is done in three levels such as total earning by each author, total advance earned by each other and royalty earned by book per each author. The use of Kimball star schema approach helped us in designing the data warehouse to extract the required information to do the analysis. The reports can be produced over different time periods to identify seasonal trend, or asses the impacts or promotions/marketing campaigns.