TASK 5: - Data Warehouse (A)

CONESTOGA COLLAGE

COURSE CODE: 1372

PROGRAM CODE- PROG8651

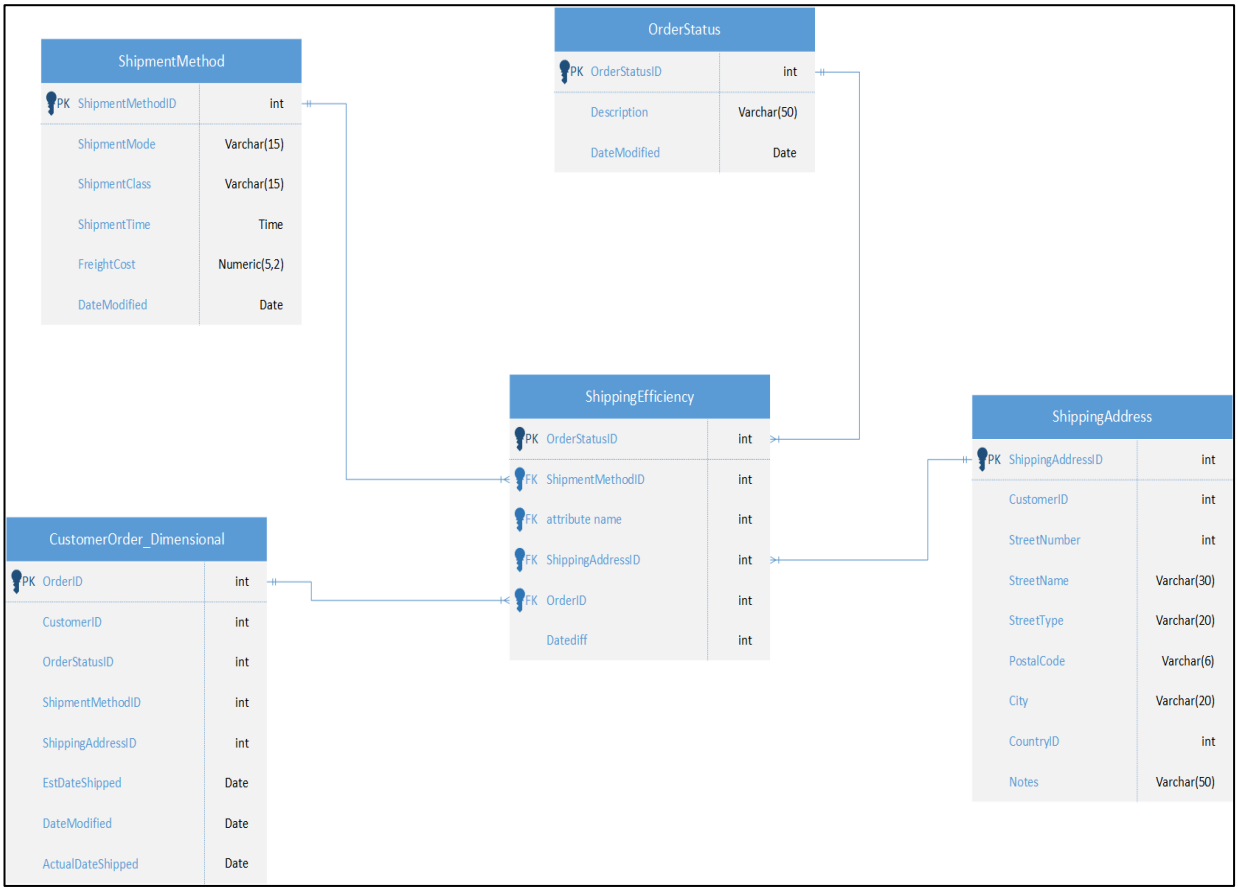
SECTION – 10

STUDENT NAME: KAUSHAL PARMAR

DATE: 17. April 2024

Q _ 1 Visio file

From: Kaushal Parmar
To: Matt Kozi
Subject: New Data Mart



Regards,

Kaushal Parmar

Q 3: - Explain why referential integrity isn't a major focus of the FACT table within a data mart based on the data warehouse star schema?

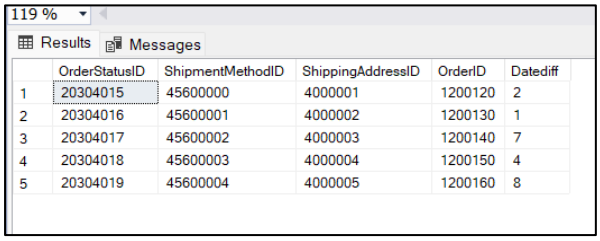
➢ What issues would arise if update and delete were enforced within the key CONSTRAINT(s)?

- The Fact table's emphasis on referential integrity is comparatively less than in traditional transactional databases when used in a data mart inside a data warehouse, mainly when working with a star schema. This purposeful method stems from the nature and goals of data warehousing, not a result of oversight. For historical reporting, fact tables in a data warehouse are commonly used to compile vast amounts of data for analysis that span time. Summaries are placed into these tables, and once loaded, they are usually stable and do not change much. More important than preserving transactional integrity or supporting operational procedures that necessitate frequent additions or deletions is enabling sophisticated inquiries and analyses, including trend analysis across time. Thus, additional complexity and performance overheads may result from guaranteeing strict referential integrity through restrictions like foreign keys. A data warehousing environment's vital query performance and data loading procedures (ETL) could be slowed down by it.

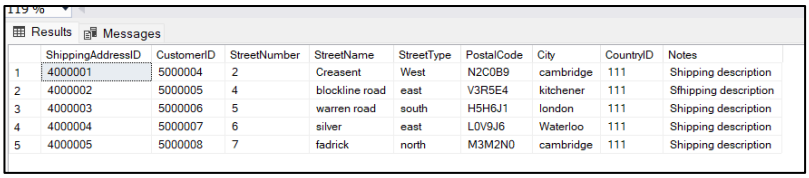
- Moreover, there could be several practical problems if changes and deletes were strictly prohibited within the essential boundaries of the Fact table. For example, removing a record from a dimension database that a fact table reference would need cascading deletes, which could lead to the removal of enormous amounts of important historical data and have disastrous effects on data analysis. Updates may also change historical accuracy, affecting the validity of trend analysis and business intelligence findings. As a result, although referential integrity is preserved to some extent to guarantee data consistency, performance and flexibility—essential for data analysis and reporting—are given precedence in a data warehouse's Fact table.

Q 4: - Sample Data

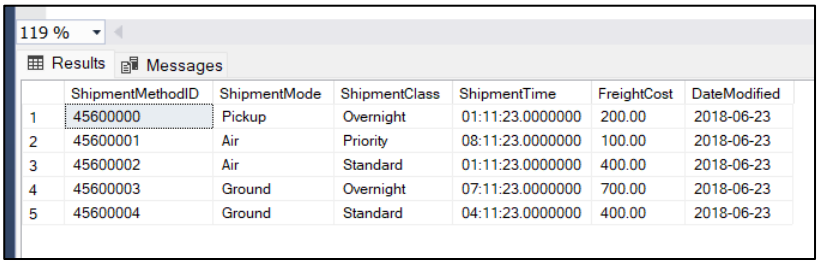
From: Kaushal Parmar
To: Matt Kozi
Subject: Sample Data
Hello, Matt
Greetings of the day. I have sent u SQL. File for Q (2) and Q (4) and you can see output pics as you said at least 5 data per table.



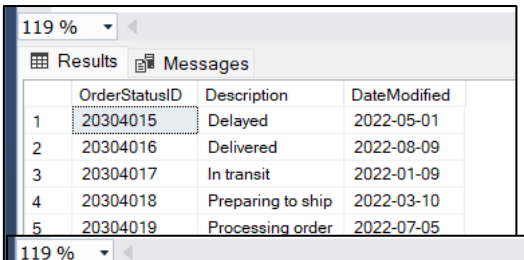
	OrderStatusID	ShipmentMethodID	ShippingAddressID	OrderID	Datediff
1	20304015	45600000	4000001	1200120	2
2	20304016	45600001	4000002	1200130	1
3	20304017	45600002	4000003	1200140	7
4	20304018	45600003	4000004	1200150	4
5	20304019	45600004	4000005	1200160	8



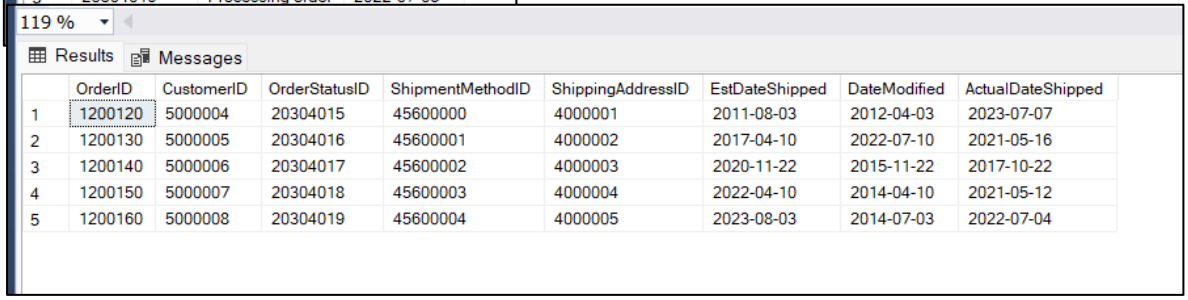
	ShippingAddressID	CustomerID	StreetNumber	StreetName	StreetType	PostalCode	City	CountryID	Notes
1	4000001	5000004	2	Creasent	West	N2C0B9	cambridge	111	Shipping description
2	4000002	5000005	4	blockline road	east	V3R5E4	kitchener	111	Shipping description
3	4000003	5000006	5	warren road	south	H5H6J1	london	111	Shipping description
4	4000004	5000007	6	silver	east	L0V9J6	Waterloo	111	Shipping description
5	4000005	5000008	7	fadrick	north	M3M2N0	cambridge	111	Shipping description



	ShipmentMethodID	ShipmentMode	ShipmentClass	ShipmentTime	FreightCost	DateModified
1	45600000	Pickup	Overnight	01:11:23.0000000	200.00	2018-06-23
2	45600001	Air	Priority	08:11:23.0000000	100.00	2018-06-23
3	45600002	Air	Standard	01:11:23.0000000	400.00	2018-06-23
4	45600003	Ground	Overnight	07:11:23.0000000	700.00	2018-06-23
5	45600004	Ground	Standard	04:11:23.0000000	400.00	2018-06-23



	OrderStatusID	Description	DateModified
1	20304015	Delayed	2022-05-01
2	20304016	Delivered	2022-08-09
3	20304017	In transit	2022-01-09
4	20304018	Preparing to ship	2022-03-10
5	20304019	Processing order	2022-07-05



	OrderID	CustomerID	OrderStatusID	ShipmentMethodID	ShippingAddressID	EstDateShipped	DateModified	ActualDateShipped
1	1200120	5000004	20304015	45600000	4000001	2011-08-03	2012-04-03	2023-07-07
2	1200130	5000005	20304016	45600001	4000002	2017-04-10	2022-07-10	2021-05-16
3	1200140	5000006	20304017	45600002	4000003	2020-11-22	2015-11-22	2017-10-22
4	1200150	5000007	20304018	45600003	4000004	2022-04-10	2014-04-10	2021-05-12
5	1200160	5000008	20304019	45600004	4000005	2023-08-03	2014-07-03	2022-07-04