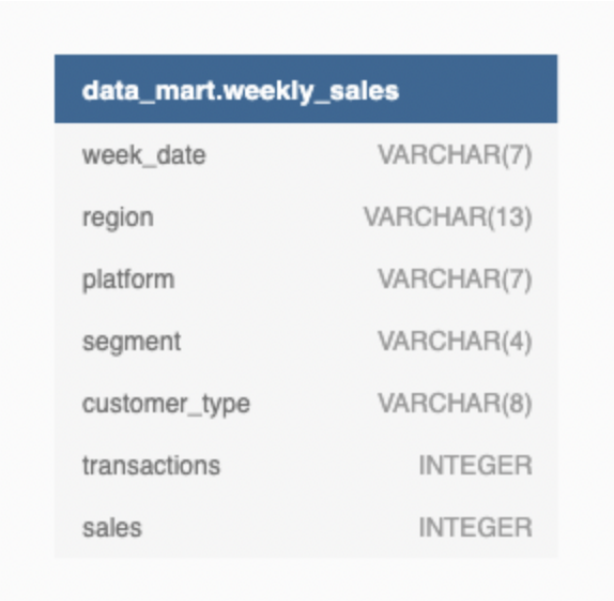
ER DIAGRAM




```sql
CREATE TABLE weekly_sales (
  "week_date" NVARCHAR(10),
  "region" NVARCHAR(20),
  "platform" NVARCHAR(20),
  "segment" NVARCHAR(4),
  "customer_type" NVARCHAR(20),
  "transactions" INTEGER,
  "sales" INTEGER
)

select * from weekly_sales;
INSERT INTO weekly_sales
  (week_date, region, platform, segment, customer_type,
transactions, sales)
VALUES
  ('31/8/20', 'ASIA', 'Retail', 'F1', 'New', '31574',
'996575'),
  ('25/12/20', 'USA', 'Retail', 'null', 'Guest', '529151',
'16509610'),
  ('01/4/20', 'INDIA', 'Retail', 'C1', 'New', '4517',
'141942'),
  ('31/10/20', 'AFRICA', 'Retail', 'C2', 'New', '58046',
'1758388'),
```

```
  ('16/3/20', 'CANADA', 'Shopify', 'F2', 'Existing',
'1336', '243878'),
  ('07/9/20', 'AFRICA', 'Shopify', 'F3', 'Existing',
'2514', '519502'),
  ('29/11/20', 'ASIA', 'Shopify', 'F1', 'Existing',
'2158', '371417'),
  ('31/8/20', 'AFRICA', 'Shopify', 'F2', 'New', '318',
'49557'),
  ('09/2/20', 'AFRICA', 'Retail', 'C3', 'New', '111032',
'3888162'),
  ('13/8/20', 'USA', 'Shopify', 'F1', 'Existing', '1398',
'260773');
```

QUESTIONS:

1. In a single query, perform the following operations and generate a new table in the data_mart

schema named clean_weekly_sales:

a. Convert the week_date to a DATE format

b. Add a week_number as the second column for each week_date value, for example any value

from the 1st of January to 7th of January will be 1, 8th to 14th will be 2 etc

c. Add a month_number with the calendar month for each week_date value as the 3rd column

d. Add a calendar_year column as the 4th column containing either 2018, 2019 or 2020 values

e. Add a new column called age_band after the original segment column using the following

mapping on the number inside the segment value

segment	age_band
1	Young Adults
2	Middle Aged

3 or 4	Retirees

f. Add a new demographic column using the following mapping for the first letter in the

segment values:

segment	demographic
C	Couples
F	Families

g. Generate a new avg_transaction column as the sales value divided by transactions rounded to 2 decimal places for each record

```
SELECT
  CONVERT(date, week_date, 3) AS week_date,
  DATEPART(week, CONVERT(date, week_date, 3)) AS
week_number,
  DATEPART(month, CONVERT(date, week_date, 3)) AS
month_number,
  DATEPART(year, CONVERT(date, week_date, 3)) AS
calendar_year,
  region,
  platform,
  segment,
  customer_type,
  CASE
    WHEN RIGHT(segment, 1) = '1' THEN 'Young Adults'
    WHEN RIGHT(segment, 1) = '2' THEN 'Middle Aged'
    WHEN RIGHT(segment, 1) IN ('3', '4') THEN 'Retirees'
    ELSE 'unknown' END AS age_band,
  CASE
    WHEN LEFT(segment, 1) = 'C' THEN 'Couples'
    WHEN LEFT(segment, 1) = 'F' THEN 'Families'
    ELSE 'unknown' END AS demographic,
  transactions,
  CAST(sales AS bigint) AS sales,
  ROUND(CAST(sales AS FLOAT)/transactions, 2) AS
avg_transaction
INTO clean_weekly_sales
FROM weekly_sales;
use casestudy;
select * from clean_weekly_sales;
```

```
    ELSE 'unknown' END AS age_band,
  CASE
    WHEN LEFT(segment, 1) = 'C' THEN 'Couples'
    WHEN LEFT(segment, 1) = 'F' THEN 'Families'
    ELSE 'unknown' END AS demographic,
  transactions,
  CAST(sales AS bigint) AS sales,
  ROUND(CAST(sales AS FLOAT)/transactions, 2) AS avg_transaction
INTO clean_weekly_sales
FROM weekly_sales;
use casestudy;
select * from clean_weekly_sales;
/*qn 2*/
SELECT
```

60 %

Results | Messages

	week_date	week_number	month_number	calendar_year	region	platform	segment	customer_type	age_band	demographic	transactions	sales	avg_transaction
1	2020-08-31	36	8	2020	ASIA	Retail	F1	New	Young Adults	Families	31574	996575	31.56
2	2020-12-25	52	12	2020	USA	Retail	null	Guest	unknown	unknown	529151	16509610	31.2
3	2020-04-01	14	4	2020	INDIA	Retail	C1	New	Young Adults	Couples	4517	141942	31.42
4	2020-10-31	44	10	2020	AFRICA	Retail	C2	New	Middle Aged	Couples	58046	1758388	30.29
5	2020-03-16	12	3	2020	CANADA	Shopify	F2	Existing	Middle Aged	Families	1336	243878	182.54
6	2020-09-07	37	9	2020	AFRICA	Shopify	F3	Existing	Retirees	Families	2514	519502	206.64
7	2020-11-29	49	11	2020	ASIA	Shopify	F1	Existing	Young Adults	Families	2158	371417	172.11
8	2020-08-31	36	8	2020	AFRICA	Shopify	F2	New	Middle Aged	Families	318	49557	155.84
9	2020-02-09	7	2	2020	AFRICA	Retail	C3	New	Retirees	Couples	111032	3888162	35.02
10	2020-08-13	33	8	2020	USA	Shopify	F1	Existing	Young Adults	Families	1398	260773	186.53

2. How many total transactions were there for each year in the dataset?

```sql
SELECT
  calendar_year,
  SUM(transactions) AS total_transactions
FROM clean_weekly_sales
GROUP BY calendar_year
ORDER BY calendar_year;
```

```
/*qn 2*/
SELECT
  calendar_year,
  SUM(transactions) AS total_transactions
FROM clean_weekly_sales
GROUP BY calendar_year
ORDER BY calendar_year;

/*qn3*/
SELECT
  region,
  month_number,
  SUM(sales) AS total_sales
FROM clean_weekly_sales
```

160 %

Results | Messages

	calendar_year	total_transactions
1	2020	742044

3. What is the total sales for each region for each month?

```sql
SELECT
  region,
```

```sql
  month_number,
  SUM(sales) AS total_sales
FROM clean_weekly_sales
GROUP BY region, month_number
ORDER BY region, month_number;
```

```sql
/*qn3*/
SELECT
  region,
  month_number,
  SUM(sales) AS total_sales
FROM clean_weekly_sales
GROUP BY region, month_number
ORDER BY region, month_number;

/*qn4*/
SELECT
```

160 %

Results | Messages

	region	month_number	total_sales
1	AFRICA	2	3888162
2	AFRICA	8	49557
3	AFRICA	9	519502
4	AFRICA	10	1758388
5	ASIA	8	996575
6	ASIA	11	371417
7	CANADA	3	243878
8	INDIA	4	141942
9	USA	8	260773
10	USA	12	16509610

4. What is the total count of transactions for each platform?

```sql
SELECT
  platform,
  SUM(transactions) AS total_transactions
FROM clean_weekly_sales
GROUP BY platform;
```

```sql
/*qn4*/
SELECT
  platform,
  SUM(transactions) AS total_transactions
FROM clean_weekly_sales
GROUP BY platform;

/*qn5*/
DECLARE @weekNum int = (
  SELECT DISTINCT week_number
  FROM clean_weekly_sales
  WHERE week_date = '2020-06-15')
```

160 %

Results | Messages

	platform	total_transactions
1	Retail	734320
2	Shopify	7724

5. What is the total sales for the 4 weeks before and after 2020-06-15?

```sql
DECLARE @weekNum int = (
  SELECT DISTINCT week_number
  FROM clean_weekly_sales
  WHERE week_date = '2020-06-15')


 SELECT
    SUM(CASE WHEN week_number BETWEEN @weekNum-4 AND
@weekNum-1 THEN sales END) AS before_changes,
    SUM(CASE WHEN week_number BETWEEN @weekNum AND
@weekNum+3 THEN sales END) AS after_changes
  FROM clean_weekly_sales
  WHERE calendar_year = 2020
```

```sql
DECLARE @weekNum int = (
  SELECT DISTINCT week_number
  FROM clean_weekly_sales
  WHERE week_date = '2020-06-15')


 SELECT
    SUM(CASE WHEN week_number BETWEEN @weekNum-4 AND @weekNum-1 THEN sales END) AS before_changes,
    SUM(CASE WHEN week_number BETWEEN @weekNum AND @weekNum+3 THEN sales END) AS after_changes
  FROM clean_weekly_sales
  WHERE calendar_year = 2020

/*qn6*/
```

160 %

Results | Messages

	before_changes	after_changes
1	NULL	NULL

6. Which areas of the business have the 5 highest negative impact in sales metrics performance in 2020 for the 12 weeks in the 2 nd quarter?

```
SELECT top 5 month_number,sum(   sales) "Sales Amount"

  FROM clean_weekly_sales
  WHERE calendar_year = 2020 and month_number>6
  group by month_number
```

```
/*qn6*/
SELECT top 5 month_number,sum(  sales) "Sales Amount"

  FROM clean_weekly_sales
  WHERE calendar_year = 2020 and month_number>6
  group by month_number



/*qn7*/
DECLARE @retailSales bigint = (
  SELECT SUM(sales)
  FROM clean_weekly_sales
```

160 %

Results | Messages

	month_number	Sales Amount
1	8	1306905
2	9	519502
3	10	1758388
4	11	371417
5	12	16509610

7. Which “age_band” and “demographic” values contribute the most to Retail sales?

```
DECLARE @retailSales bigint = (
  SELECT SUM(sales)
  FROM clean_weekly_sales
  WHERE platform = 'Retail')

SELECT
  age_band,
  demographic,
  SUM(sales) AS sales,
  CAST(100.0 * SUM(sales)/@retailSales AS decimal(5, 2))
AS contribution
FROM clean_weekly_sales
WHERE platform = 'Retail'
GROUP BY age_band, demographic
ORDER BY contribution DESC;
```

```sql
DECLARE @retailSales bigint = (
  SELECT SUM(sales)
  FROM clean_weekly_sales
  WHERE platform = 'Retail')

SELECT
  age_band,
  demographic,
  SUM(sales) AS sales,
  CAST(100.0 * SUM(sales)/@retailSales AS decimal(5, 2)) AS contribution
FROM clean_weekly_sales
WHERE platform = 'Retail'
GROUP BY age_band, demographic
```

160 %

Results | Messages

	age_band	demographic	sales	contribution
1	unknown	unknown	16509610	70.87
2	Retirees	Couples	3888162	16.69
3	Middle Aged	Couples	1758388	7.55
4	Young Adults	Families	996575	4.28
5	Young Adults	Couples	141942	0.61