SQL
DATA–ANALYSIS

of

Project: Providing Tourism Insights to the Telangana Government.

(A Codebasics Resume Challenge Project)

Module–1

INTRODUCTION

THIS IS MODULE 1 OF A DATA ANALYSIS PROJECT FOCUSED ON PROVIDING VALUABLE INSIGHTS TO THE TELANGANA TOURISM DEPARTMENT.

USING SQL ANALYSIS AND DATA CLEANING TECHNIQUES, I HAVE EXTRACTED KEY INSIGHTS AND ADDRESSED RESEARCH QUESTIONS. STAY TUNED FOR MORE MODULES TO EXPLORE FURTHER INSIGHTS AND RECOMMENDATIONS.

[QUESTION-1]
LIST DOWN THE TOP 10 DISTRICTS THAT HAVE THE HIGHEST NUMBER OF DOMESTIC VISITORS OVERALL (2016-2019)?

```sql
-- [Query-1]: Top 10 Districts (Domestic)

select  district,
        sum(visitors) as overall_visitors
from telangana_project.modify_domestic_visitors
group by district
order by overall_visitors desc limit 10;
```

district	overall_visitors
Hyderabad	83900960
Rajanna Sircilla	41763276
Warangal (Urban)	30726603
Yadadri Bhongir	26893080
Bhadradri Kothagudem	21600962
Medak	20542639
Jayashankar Bhoopalpally	19632865
Mahbubnagar	17180118
Nirmal	13315796
Jagtial	11303514

[QUESTION-2]

LIST DOWN THE TOP 3 DISTRICTS BASED ON COMPOUND ANNUAL GROWTH RATE (CAGR) OF VISITORS BETWEEN (2016-2019)?

```sql
-- [Query-1]: For Domestic

with cte as (
select  district,
        sum(case when year=2016 then visitors else 0 end) as Initial_Value,
        sum(case when year=2019 then visitors else 0 end) as Final_Value
from telangana_project.modify_domestic_visitors
group by district
)
select *, round((power(Final_Value/Initial_Value,1/3)-1)*100,2) as CAGR
from cte
order by CAGR desc limit 5;
```

district	Initial_Value	Final_Value	CAGR
Mancherial	7802	269810	225.8
Warangal (Rural)	19400	353500	163.15
Bhadradri Kothagudem	889030	12817737	143.39

[QUESTION-3]

LIST DOWN THE BOTTOM 3 DISTRICTS BASED ON COMPOUNDED ANNUAL GROWTH RATE(CAGR) OF VISITORS BETWEEN (2016-2019)?

```sql
-- For Domestic
with cte as (
select district,
        sum(case when year=2016 then visitors else 0 end) as Initial_Value,
        sum(case when year=2019 then visitors else 0 end) as Final_Value
from telangana_project.modify_domestic_visitors
group by district
)
select *, round((power(Final_Value/Initial_Value,1/3)-1)*100,2) as CAGR
from cte
where Initial_Value != 0
order by CAGR asc limit 3;
```

district	Initial_Value	Final_Value	CAGR
Karimnagar	9167468	77491	-79.63
Nalgonda	5858461	140918	-71.13
Warangal (Urban)	25788035	1795230	-58.86

[QUESTION-4]

WHAT ARE THE PEAK AND LOW SEASON MONTHS FOR HYDERABAD BASED ON THE DATA FROM 2016 TO 2019 FOR HYDERABAD DISTRICT?

```
-- [Query-1] For Domestic (Top-3 peak_month)

select  month as peak_month,
  sum(visitors) as visitors
from telangana_project.modify_domestic_visitors
where district="hyderabad"
group by month
order by visitors desc limit 3;
```

```
-- [Query-2] For Domestic (bottom-3 low_month )

select  month as low_month,
  sum(visitors) as visitors
from telangana_project.modify_domestic_visitors
where district="Hyderabad"
group by month
order by visitors asc limit 3;
```

peak_month	visitors
June	16897783
December	9338637
October	6552397

low_month	visitors
February	5014430
March	5227626
September	5312283

[QUESTION-5]

SHOW THE TOP & BOTTOM 3 DISTRICTS WITH HIGH DOMESTIC TO FOREIGN TOURIST RATIO?

```sql
with cte as (
select d.district,
       sum(d.visitors) as d_visitors,
       sum(f.visitors) as f_visitors
from telangana_project.modify_domestic_visitors d
join telangana_project.modify_foreign_visitors f
on d.district=f.district and d.year=f.year and d.month=f.month
group by district
),
cte2 as (
select *, round(d_visitors/f_visitors,0) as ratio
from cte
where f_visitors ≠ 0
order by ratio
),
cte3 as (
select *, rank() over(order by ratio) as top,
          rank() over(order by ratio desc) as bottom
from cte2
)
select district, d_visitors, f_visitors, ratio, top
from cte3
where top<4 or bottom<4
```

district	d_visitors	f_visitors	ratio	top
Hyderabad	83900960	1044898	80	1
Warangal (Rural)	819162	306	2677	2
Mulugu	1819800	575	3165	3
Adilabad	7321575	32	228799	11
Jangaon	826280	2	413140	12
Nirmal	13315796	2	6657898	13

[QUESTION-6]

LIST THE TOP AND BOTTOM 5 DISTRICTS BASED ON 'POPULATION TO TOURIST FOOTFALL RATIO'* RATIO IN 2019?(*RATIO: TOTAL VISITORS/TOTAL RESIDENTS POPULATION IN THE GIVEN YEAR)

```sql
with cte as (
select D.district, sum(D.visitors) as d_visitors,
            sum(F.visitors) as f_visitors,
            sum(D.visitors)+sum(F.visitors) as t_visitors
from telangana_project.modify_domestic_visitors D
join telangana_project.modify_foreign_visitors F
on F.district=D.district and F.month=D.month and F.year=D.year
where D.year=2019
group by district
),
cte1 as (
select C.*, population_2019 as t_residents,
        round(t_visitors/population_2019,2) as footfall_ratio
from cte C
join population_year_wise Y
on Y.district=C.district
where t_visitors ≠ 0
)
select district, footfall_ratio, top
from ( select *, row_number() over(order by footfall_ratio desc) as top,
                row_number() over(order by footfall_ratio asc) as bottom
    from cte1) as subquery
where top<6 or bottom<6
order by footfall_ratio desc;
```

district	footfall_ratio	top
Rajanna Sircilla	28.66	1
Bhadradri Kothagudem	11.27	2
Medak	6.68	3
Mulugu	6.64	4
Yadadri Bhongir	5.71	5
Karimnagar	0.07	24
Komaram Bheem Asifabad	0.03	25
Nizamabad	0.03	26
Peddapalli	0.02	27
Kamareddy	0.00	28

[QUESTION:7]

WHAT WILL BE THE PROJECTED NUMBER OF DOMESTIC AND FOREIGN TOURISTS IN HYDERABAD IN 2025 BASED ON THE GROWTH RATE FROM PREVIOUS YEARS?

```sql
-- [Query:-1] Domestic Projected Value for 2025

with cte as (
select district, sum(case when year=2016 then visitors else 0 end) as Initial_Value,
    sum(case when year=2019 then visitors else 0 end) as Final_Value
from telangana_project.modify_domestic_visitors
where district="Hyderabad"
)
select *, round(Final_Value*power((1+AGR),6),0) as projected_dom_2025
from (
  select *, round(power(Final_Value/Initial_Value, (1/3))-1,2) as AGR     # AGR=-0.16
  from cte
    ) as subquery
```

district	Initial_Value	Final_Value	AGR	projected_dom_2025
Hyderabad	23394705	13802362	-0.16	4848743

[QUESTION-8]

ESTIMATED THE PROJECTED REVENUE FOR HYDERABAD IN 2025 BASED ON AVERAGE SPEND PER TOURIST (APPROXIMATE DATA)

TOURIST | AVERAGE REVENUE
FOREIGN | RS. 5600.00
DOMESTIC | RS. 1200.00

```sql
-- [Query-1] For Domestic Estimation

with cte as (
select district, sum(case when year=2016 then visitors else 0 end) as Initial_Value,
                sum(case when year=2019 then visitors else 0 end) as Final_Value
from telangana_project.modify_domestic_visitors
where district="hyderabad"
)
select *, round(Final_Value*power((1+AGR),6),0)  as  projected_dom_2025,
        round(Final_Value*power((1+AGR),6),0)*1200  as  Projected_dom_revenue
from (
    select *, round(power(Final_Value/Initial_Value, (1/3))-1,2) as AGR      # AGR=-0.16
    from cte
  ) as subquery
```

district	Initial_Value	Final_Value	AGR	projected_dom_2025	Projected_dom_revenue
Hyderabad	23394705	13802362	-0.16	4848743	5818491600

THANK YOU !

FOLLOW ME ON LINKEDIN FOR REGULAR UPDATES