World Database SQL Assignment

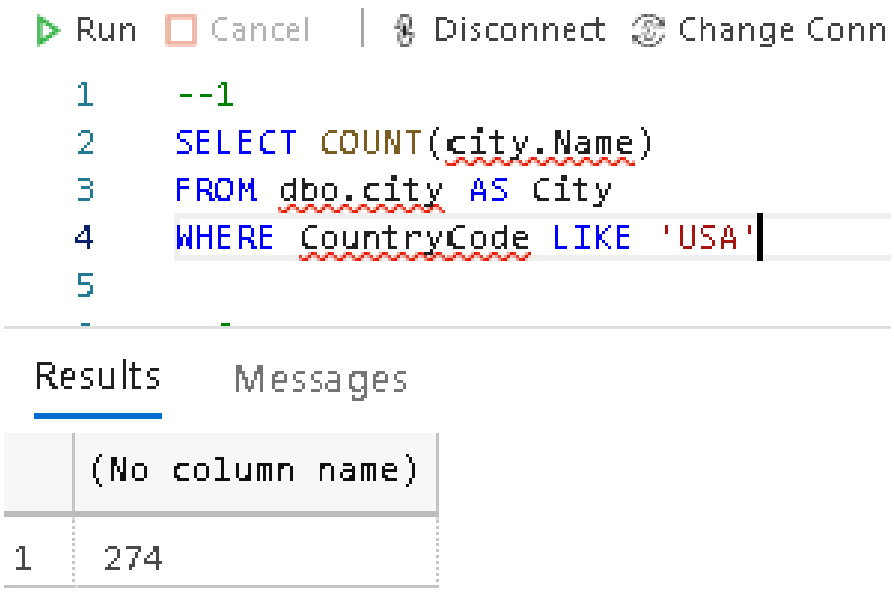
By Emma Kaas Andersen

Submission Date: 10/05/2024

Image: Accelebrate (2022)

Tasks

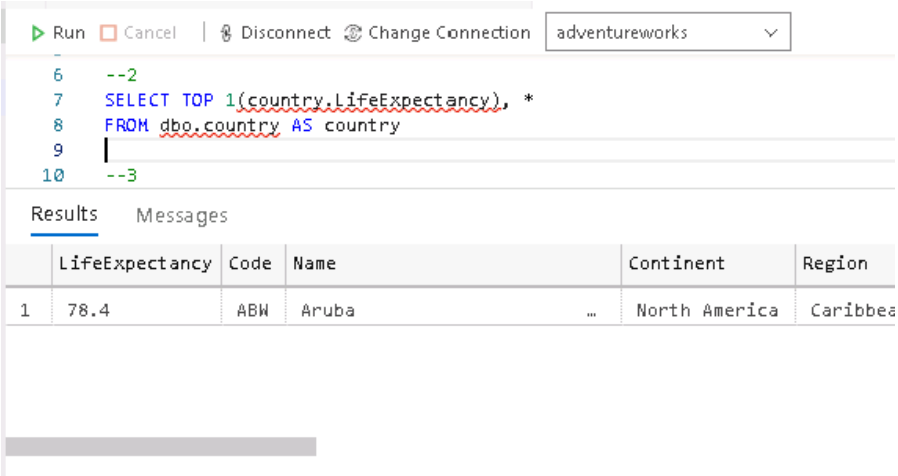
1. **Count Cities in USA:** *Scenario:* You've been tasked with conducting a demographic analysis of cities in the United States. Your first step is to determine the total number of cities within the country to provide a baseline for further analysis.



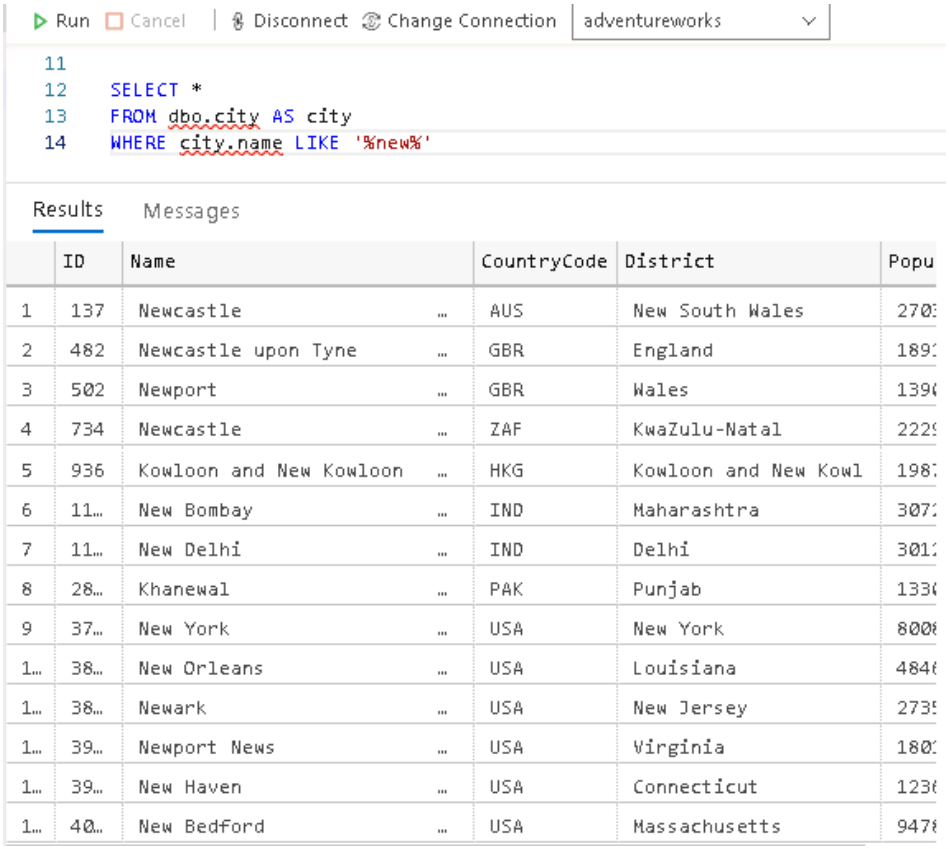
I used the COUNT() function in order to count the number of cities and then using the WHERE function was able to specify only to return cities who had the country code 'USA'.

2. **Country with Highest Life Expectancy:** *Scenario:* As part of a global health initiative, you've been assigned to identify the country with the highest life expectancy. This information will be crucial for prioritizing healthcare resources and interventions.

I used the TOP() function to find highest values in the life expectancy column from the country table which returned Aruba.



3. **"New Year Promotion: Featuring Cities with 'New :** *Scenario:* In anticipation of the upcoming New Year, your travel agency is gearing up for a special promotion featuring cities with names including the word 'New'. You're tasked with swiftly compiling a list of all cities from around the world. This curated selection will be essential in creating promotional materials and enticing travellers with exciting destinations to kick off the New Year in style.

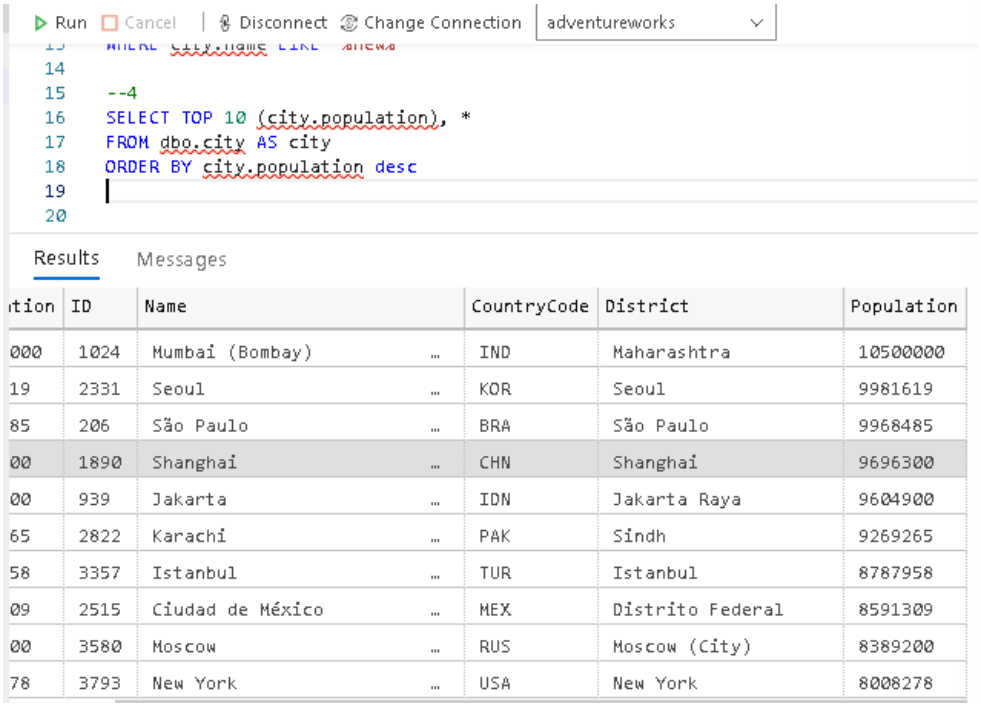


	ID	Name	CountryCode	District	Popu
1	137	Newcastle …	AUS	New South Wales	270
2	482	Newcastle upon Tyne …	GBR	England	189
3	502	Newport …	GBR	Wales	139
4	734	Newcastle …	ZAF	KwaZulu-Natal	222
5	936	Kowloon and New Kowloon …	HKG	Kowloon and New Kowl	198
6	11…	New Bombay …	IND	Maharashtra	307
7	11…	New Delhi …	IND	Delhi	301
8	28…	Khanewal …	PAK	Punjab	133
9	37…	New York …	USA	New York	800
1…	38…	New Orleans …	USA	Louisiana	484
1…	38…	Newark …	USA	New Jersey	273
1…	39…	Newport News …	USA	Virginia	180
1…	39…	New Haven …	USA	Connecticut	123
1…	40…	New Bedford …	USA	Massachusetts	947

To find the cities which include the word ‘new’ I used the wildcard character linked with the LIKE operator. Adding % means that any characters can be there which is why I used ‘%new%’. This asks SQL to find any city names which include the word ‘new’ with any characters before and after those letters together.

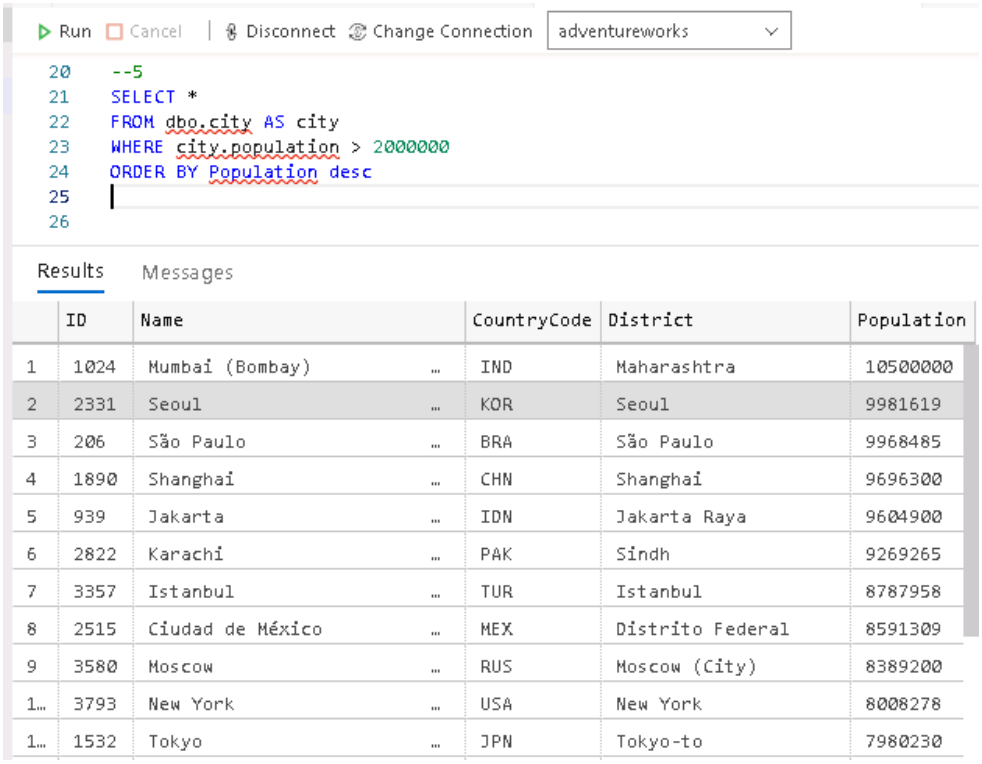
4. **Display Columns with Limit (First 10 Rows):** *Scenario:* You're tasked with providing a brief overview of the most populous cities in the world. To keep the report concise, you're instructed to list only the first 10 cities by population from the database.

I used the ORDER BY keyword to make sure that the table was ordered according to the population column and added ‘desc’ for descending order (going from highest to lowest). From there in my SELECT statement I used TOP() to only show the top 10 results



ıtion	ID	Name	CountryCode	District	Population
000	1024	Mumbai (Bombay) …	IND	Maharashtra	10500000
19	2331	Seoul …	KOR	Seoul	9981619
85	206	São Paulo …	BRA	São Paulo	9968485
00	1890	Shanghai …	CHN	Shanghai	9696300
00	939	Jakarta …	IDN	Jakarta Raya	9604900
65	2822	Karachi …	PAK	Sindh	9269265
58	3357	Istanbul …	TUR	Istanbul	8787958
09	2515	Ciudad de México …	MEX	Distrito Federal	8591309
00	3580	Moscow …	RUS	Moscow (City)	8389200
78	3793	New York …	USA	New York	8008278

5. **Cities with Population Larger than 2,000,000:** *Scenario:* A real estate developer is interested in cities with substantial population sizes for potential investment opportunities. You're tasked with identifying cities from the database with populations exceeding 2 million to focus their research efforts.



	ID	Name	…	CountryCode	District	Population
1	1024	Mumbai (Bombay)	…	IND	Maharashtra	10500000
2	2331	Seoul	…	KOR	Seoul	9981619
3	206	São Paulo	…	BRA	São Paulo	9968485
4	1890	Shanghai	…	CHN	Shanghai	9696300
5	939	Jakarta	…	IDN	Jakarta Raya	9604900
6	2822	Karachi	…	PAK	Sindh	9269265
7	3357	Istanbul	…	TUR	Istanbul	8787958
8	2515	Ciudad de México	…	MEX	Distrito Federal	8591309
9	3580	Moscow	…	RUS	Moscow (City)	8389200
1…	3793	New York	…	USA	New York	8008278
1…	1532	Tokyo	…	JPN	Tokyo-to	7980230

So that SQL only returned results for cities which had a population over 2,000,000, I used the WHERE clause including the greater than (>) symbol.

6. **Cities Beginning with 'Be' Prefix:** *Scenario:* A travel blogger is planning a series of articles featuring cities with unique names. You're tasked with compiling a list of cities from the database that start with the prefix 'Be' to assist in the blogger's content creation process.

I used the wildcard character again here much like task 3, only this time I only added the '%' symbol after 'be' so that SQL would only return results which start with 'be' and can have any characters after but not before.

Run | Cancel | Disconnect | Change Connection | adventureworks

```
24  ORDER BY Population desc
25
26  --6
27  SELECT *
28  FROM dbo.city as city
29  WHERE City.name LIKE 'be%'
30
```

Results | Messages

	ID	Name		CountryCode	District	Population
1	59	Benguela	…	AGO	Benguela	128300
2	93	Berazategui	…	ARG	Buenos Aires	276916
3	184	Belize City	…	BLZ	Belize City	55810
4	185	Belmopan	…	BLZ	Cayo	7105
5	209	Belo Horizonte	…	BRA	Minas Gerais	2139125
6	216	Belém	…	BRA	Pará	1186926
7	246	Belford Roxo	…	BRA	Rio de Janeiro	425194
8	266	Betim	…	BRA	Minas Gerais	302108
9	453	Bento Gonçalves	…	BRA	Rio Grande do Sul	89254
1…	469	Belfast	…	GBR	North Ireland	287500
1…	724	Benoni	…	ZAF	Gauteng	365467
1…	949	Bekasi	…	IDN	West Java	644300

7. **Cities with Population Between 500,000-1,000,000:** *Scenario:* An urban planning committee needs to identify mid-sized cities suitable for infrastructure development projects. You're tasked with identifying cities with populations ranging between 500,000 and 1 million to inform their decision-making process.

Run | Cancel | Disconnect | Change Connection | adventureworks

```
29  WHERE City.name LIKE 'be%'
30
31  --7
32  SELECT *
33  FROM dbo.city AS city
34  WHERE city.population <1000000 AND city.population >500000
35  ORDER BY city.population desc
36
```

Results | Messages

	ID	Name		CountryCode	District	Population
1	3214	Mogadishu	…	SOM	Banaadir	997000
2	3592	Volgograd	…	RUS	Volgograd	993400
3	1543	Sendai	…	JPN	Miyagi	989975
4	2829	Peshawar	…	PAK	Nothwest Border Prov	988005
5	1925	Baotou	…	CHN	Inner Mongolia	980000
6	134	Adelaide	…	AUS	South Australia	978100
7	1042	Madurai	…	IND	Tamil Nadu	977856
8	3175	Mekka	…	SAU	Mekka	965700
9	3071	Köln	…	DEU	Nordrhein-Westfalen	962507
10	2734	Managua	…	NIC	Managua	959000
11	3802	Detroit	…	USA	Michigan	951270
12	1926	Shenzhen	…	CHN	Guangdong	950500

The WHERE clause I used here has 2 conditions separated by AND, meaning that all results must fulfil both requirements. Then I used the greater than (>) and less than (<) symbols to specify the range of number which can be returned.

8. **Display Cities Sorted by Name in Ascending Order:** *Scenario:* A geography teacher is preparing a lesson on alphabetical order using city names. You're tasked with providing a sorted list of cities from the database in ascending order by name to support the lesson plan.

I used the ORDER BY keyword to order the city names alphabetically. SQL automatically sorts in ascending order, there is only need to specify 'desc' if the list needed to be in reverse alphabetical.

```
36
37    --8
38    SELECT *
39    FROM dbo.city AS city
40    ORDER BY city.name
```

Results Messages

	ID	Name	CountryCode	District	Population
1	698	[San Cristóbal de] la Lagu…	ESP	Canary Islands	127945
2	3665	Šahty …	RUS	Rostov-na-Donu	221800
3	2450	Šiauliai …	LTU	Šiauliai	146563
4	3479	Šostka …	UKR	Sumy	90000
5	3736	Štšolkovo …	RUS	Moskova	104900
6	548	Šumen …	BGR	Varna	94686
7	20	´s-Hertogenbosch …	NLD	Noord-Brabant	129170
8	670	A Coruña (La Coruña) …	ESP	Galicia	243402
9	3097	Aachen …	DEU	Nordrhein-Westfalen	243825
10	3318	Aalborg …	DNK	Nordjylland	161161
11	2760	Aba …	NGA	Imo & Abia	298900
12	1404	Abadan …	IRN	Khuzestan	206073

9. **Most Populated City:** *Scenario:* A real estate investment firm is interested in cities with significant population densities for potential development projects. You're tasked with identifying the most populated city from the database to guide their investment decisions and strategic planning.

Run Cancel Disconnect Change Connection adventureworks

```
39    FROM dbo.city AS city
40    ORDER BY city.name
41
42    --9
43    SELECT TOP 1 (city.population), *
44    FROM dbo.city AS city
45
```

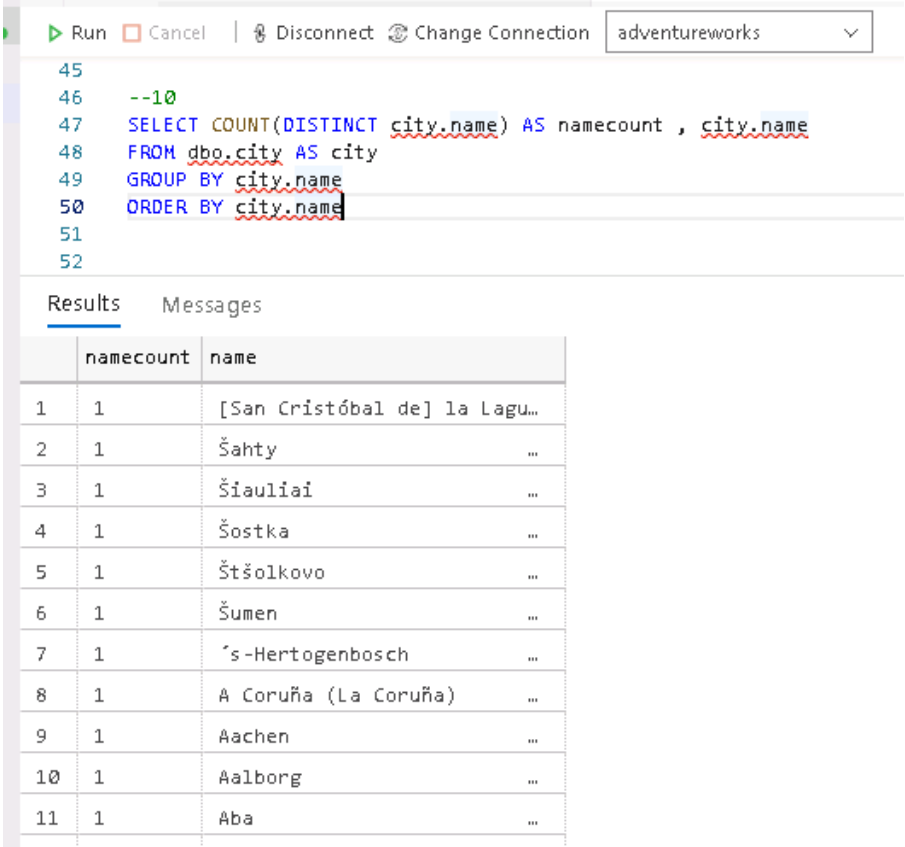
Results Messages

	population	ID	Name	CountryCode	District	Pop
1	1780000	1	Kabul …	AFG	Kabol	17

To find the city with the highest population I used TOP() which came back with Kabul.

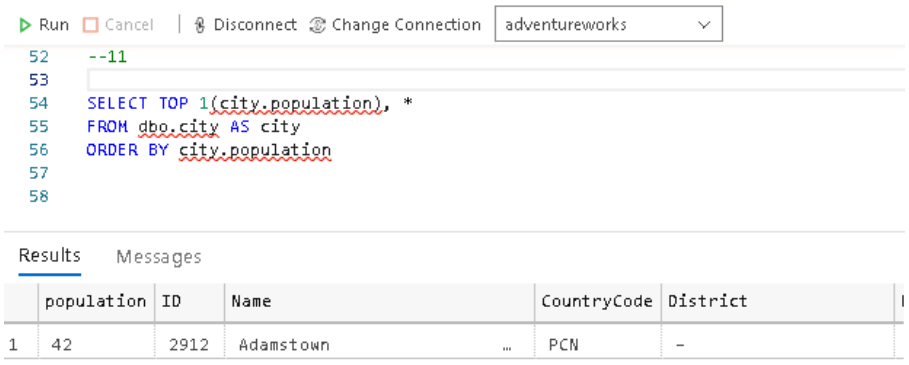
10. **City Name Frequency Analysis: Supporting Geography Education** *Scenario*: In a geography class, students are learning about the distribution of city names around the world. The teacher, in preparation for a lesson on city name frequencies, wants to provide students with a list of unique city names sorted alphabetically, along with their respective counts of occurrences in the database. You're tasked with this sorted list to support the geography teacher's l

To find how many time each city name is used, I used the COUNT() function along with distinct and made a new column called name count. This will return how many times each city name was listed under the city name column.



	namecount	name
1	1	[San Cristóbal de] la Lagu…
2	1	Šahty …
3	1	Šiauliai …
4	1	Šostka …
5	1	Štšolkovo …
6	1	Šumen …
7	1	´s-Hertogenbosch …
8	1	A Coruña (La Coruña) …
9	1	Aachen …
10	1	Aalborg …
11	1	Aba …

11. **City with the Lowest Population:** *Scenario:* A census bureau is conducting an analysis of urban population distribution. You're tasked with identifying the city with the lowest population from the database to provide a comprehensive overview of demographic trends.



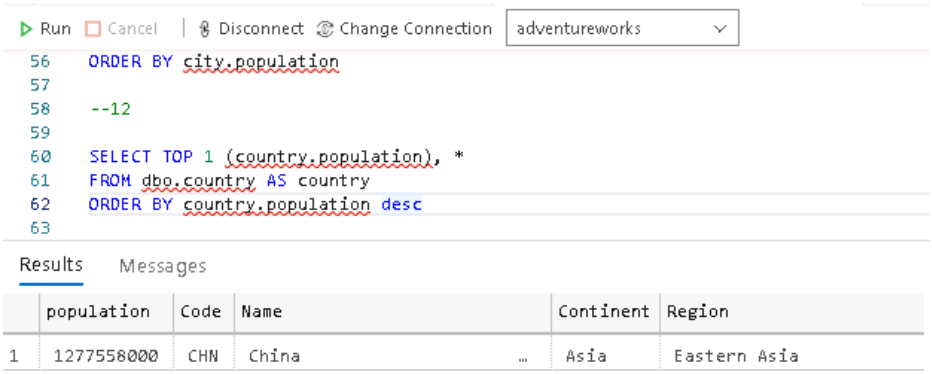
	population	ID	Name	CountryCode	District
1	42	2912	Adamstown …	PCN	-

I first ordered the table by population so that the lowest population would be at the top. Then used TOP() to only return the first result.

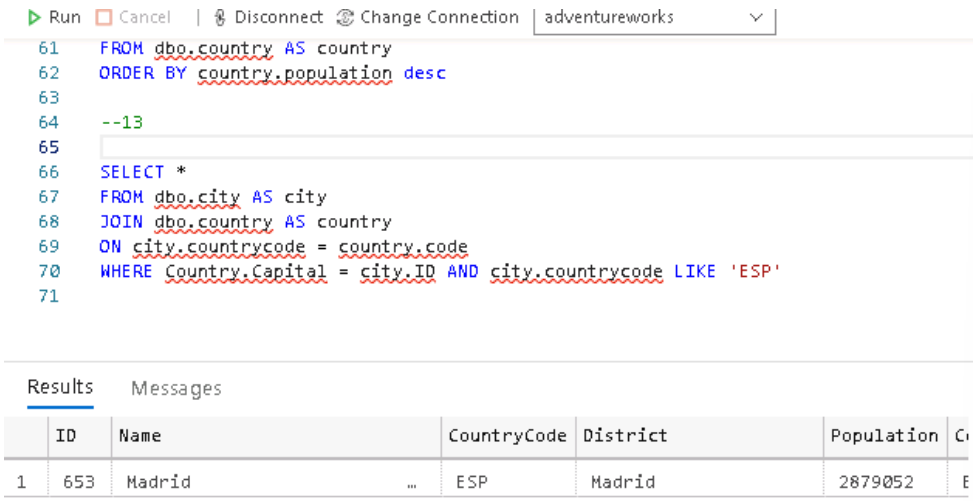
12. **Country with Largest Population:** *Scenario:* A global economic research institute requires data on countries with the largest populations for a comprehensive analysis. You're tasked with identifying

the country with the highest population from the database to provide valuable insights into demographic trends.

This query is similar to the previous, however here I queried the country table and ordered by population in descending order so that the largest value would be the first result.



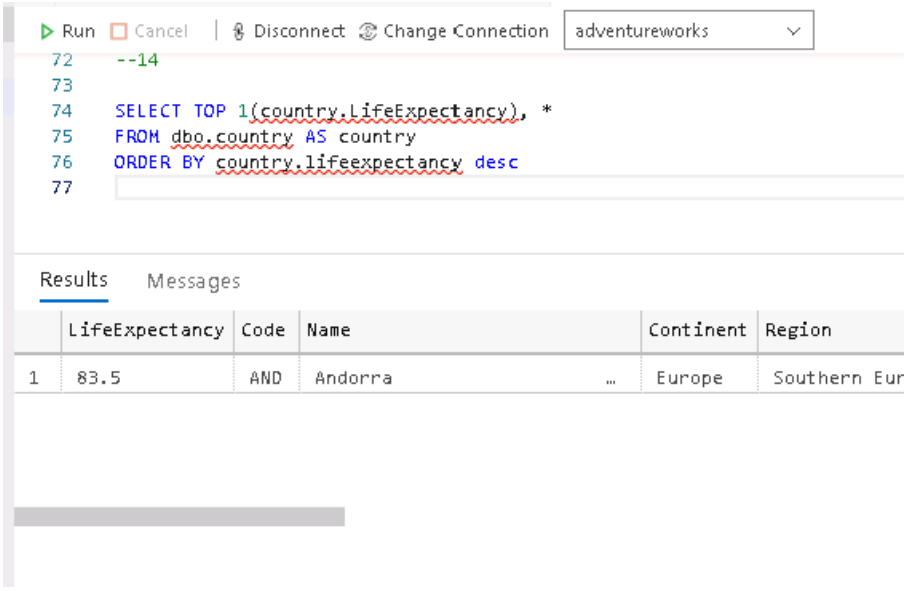
13. **Capital of Spain:** *Scenario:* A travel agency is organizing tours across Europe and needs accurate information on capital cities. You're tasked with identifying the capital of Spain from the database to ensure itinerary accuracy and provide travellers with essential destination information.



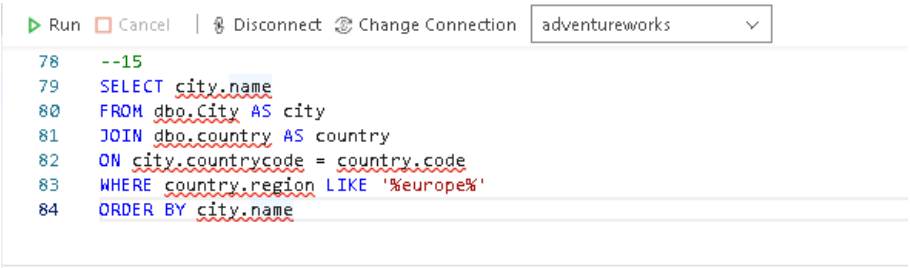
I used the JOIN clause to join the city and country tables. Then I used WHERE to specify that I only want to return results from the city table which matches the capital ID column in the country table, and also have the country code 'ESP' to only return the capital of Spain.

14. **Country with Highest Life Expectancy:** *Scenario:* A healthcare foundation is conducting research on global health indicators. You're tasked with identifying the country with the highest life expectancy from the database to inform their efforts in improving healthcare systems and policies.

I used ORDER BY to order the table by life expectancy descending, and then in SELECT specified only to return the first result.



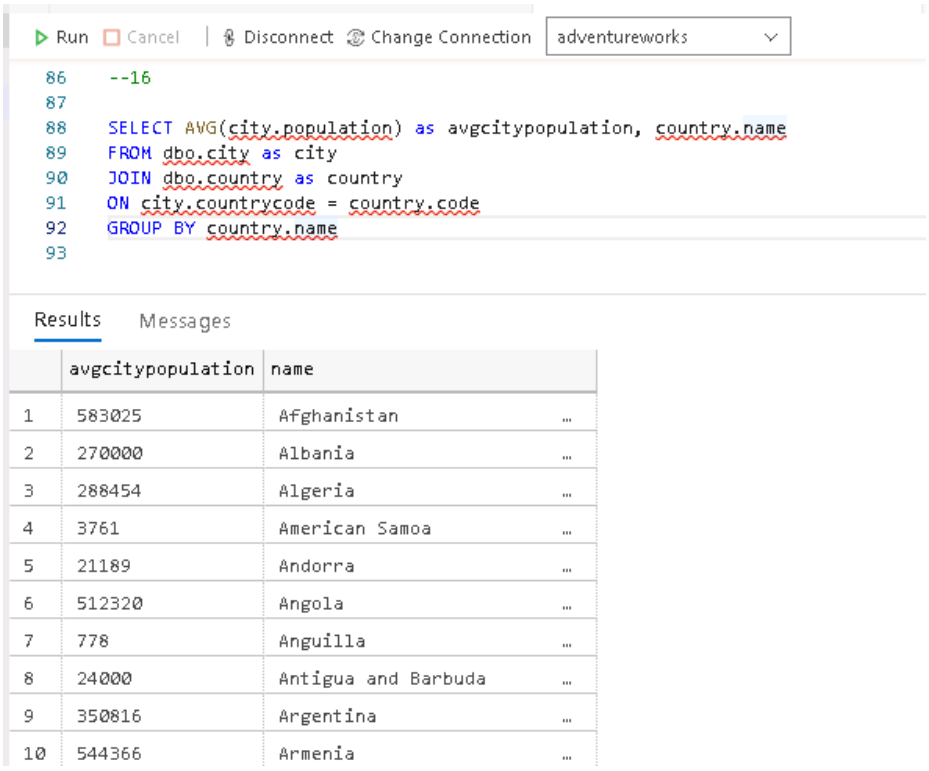
15. **Cities in Europe:** *Scenario:* A European cultural exchange program is seeking to connect students with cities across the continent. You're tasked with compiling a list of cities located in Europe from the database to facilitate program planning and student engagement.



I used JOIN to connect the city and country tables as I need to refer to columns in both. Then using WHERE I specified that I only wanted results where the region included ‘europe’ this way when I SELECT the city names it will only result in cities from Europe.

16. **Average Population by Country:** *Scenario:* A demographic research team is conducting a comparative analysis of population distributions across countries. You're tasked with calculating the average population for each country from the database to provide valuable insights into global population trends.

I used JOIN to join the city and country tables. Then I used the AVG() function in the SELECT statement to average city populations and GROUP BY to group each of those averages in their respective countries.

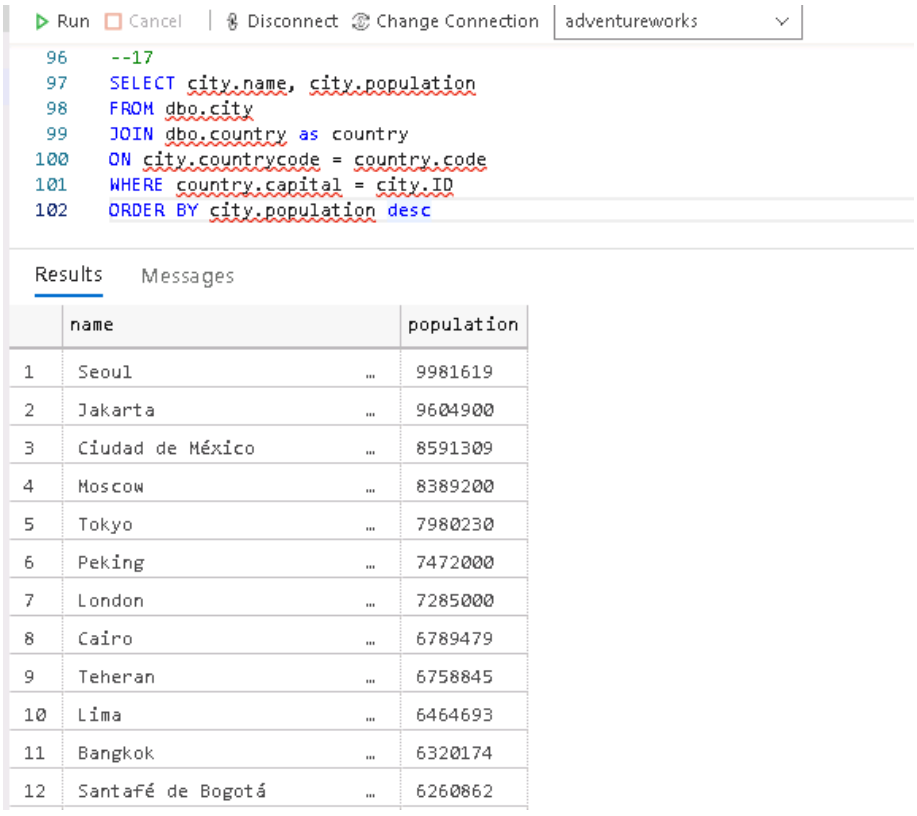


```
Run | Cancel | Disconnect | Change Connection | adventureworks

86  --16
87
88  SELECT AVG(city.population) as avgcitypopulation, country.name
89  FROM dbo.city as city
90  JOIN dbo.country as country
91  ON city.countrycode = country.code
92  GROUP BY country.name
93
```

Results | Messages

	avgcitypopulation	name	…
1	583025	Afghanistan	…
2	270000	Albania	…
3	288454	Algeria	…
4	3761	American Samoa	…
5	21189	Andorra	…
6	512320	Angola	…
7	778	Anguilla	…
8	24000	Antigua and Barbuda	…
9	350816	Argentina	…
10	544366	Armenia	…

17. **Capital Cities Population Comparison:** *Scenario:* A statistical analysis firm is examining population distributions between capital cities worldwide. You're tasked with comparing the populations of capital cities from different countries to identify trends and patterns in urban demographics.



```
Run | Cancel | Disconnect | Change Connection | adventureworks

96  --17
97  SELECT city.name, city.population
98  FROM dbo.city
99  JOIN dbo.country as country
100 ON city.countrycode = country.code
101 WHERE country.capital = city.ID
102 ORDER BY city.population desc
```

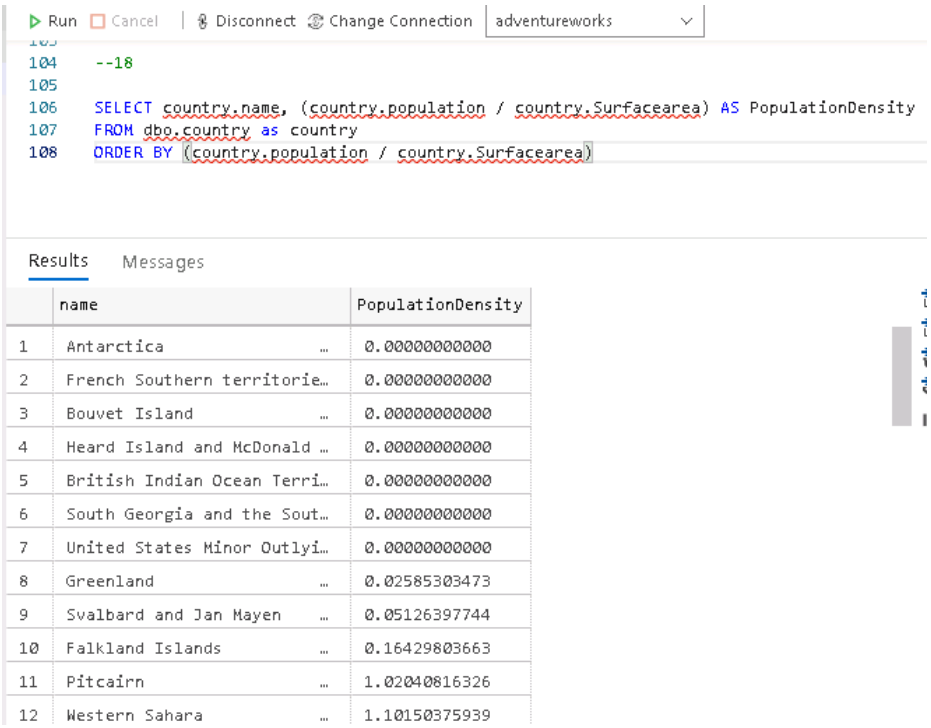
Results | Messages

	name	…	population
1	Seoul	…	9981619
2	Jakarta	…	9604900
3	Ciudad de México	…	8591309
4	Moscow	…	8389200
5	Tokyo	…	7980230
6	Peking	…	7472000
7	London	…	7285000
8	Cairo	…	6789479
9	Teheran	…	6758845
10	Lima	…	6464693
11	Bangkok	…	6320174
12	Santafé de Bogotá	…	6260862

After joining the city and country tables, I used WHERE to specify that I only wanted results where the country table's capital number ID is equal (=) to the ID in the city table. This was it would exclude any cities which don't have matching ID's to the capital cities.

18. **Countries with Low Population Density:** *Scenario:* An agricultural research institute is studying countries with low population densities for potential agricultural development projects. You're tasked with identifying countries with sparse populations from the database to support the institute's research efforts.

To find the population density I divided the country population by the surface area. Then I used ORDER By so that the table was ordered by population density in ascending order (from lowest to highest).

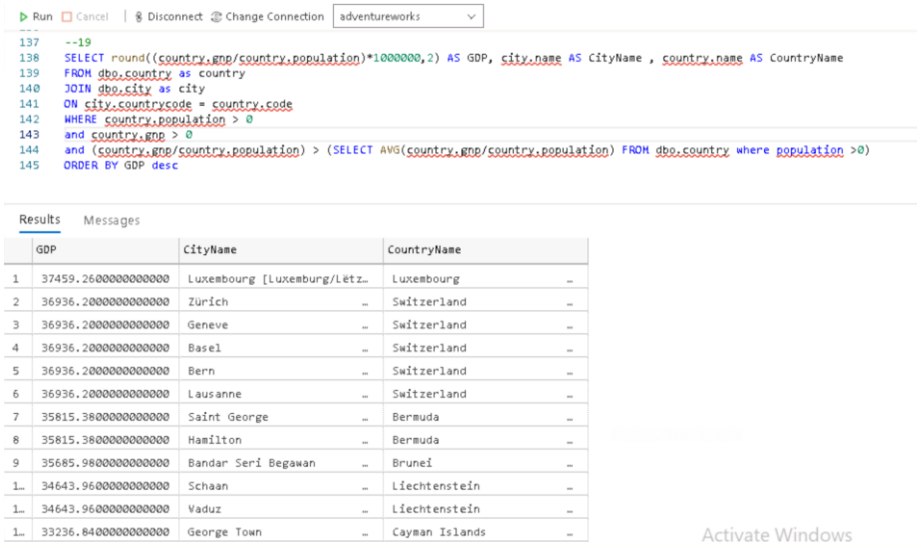


```
Run | Cancel | Disconnect | Change Connection | adventureworks

--18

SELECT country.name, (country.population / country.Surfacearea) AS PopulationDensity
FROM dbo.country as country
ORDER BY (country.population / country.Surfacearea)
```

Results | Messages

	name	PopulationDensity
1	Antarctica	0.00000000000
2	French Southern territorie…	0.00000000000
3	Bouvet Island	0.00000000000
4	Heard Island and McDonald …	0.00000000000
5	British Indian Ocean Terri…	0.00000000000
6	South Georgia and the Sout…	0.00000000000
7	United States Minor Outlyi…	0.00000000000
8	Greenland	0.02585303473
9	Svalbard and Jan Mayen	0.05126397744
10	Falkland Islands	0.16429803663
11	Pitcairn	1.02040816326
12	Western Sahara	1.10150375939

19. **Cities with High GDP per Capita:** *Scenario:* An economic consulting firm is analyzing cities with high GDP per capita for investment opportunities. You're tasked with identifying cities with above-average GDP per capita from the database to assist the firm in identifying potential investment destinations.



```
Run | Cancel | Disconnect | Change Connection | adventureworks

--19
SELECT round((country.gnp/country.population)*1000000,2) AS GDP, city.name AS CityName , country.name AS CountryName
FROM dbo.country as country
JOIN dbo.city as city
ON city.countrycode = country.code
WHERE country.population > 0
and country.gnp > 0
and (country.gnp/country.population) > (SELECT AVG(country.gnp/country.population) FROM dbo.country where population >0)
ORDER BY GDP desc
```

Results | Messages

	GDP	CityName	CountryName
1	37459.2600000000000	Luxembourg [Luxemburg/Lëtz…	Luxembourg
2	36936.2000000000000	Zürich	Switzerland
3	36936.2000000000000	Geneve	Switzerland
4	36936.2000000000000	Basel	Switzerland
5	36936.2000000000000	Bern	Switzerland
6	36936.2000000000000	Lausanne	Switzerland
7	35815.3800000000000	Saint George	Bermuda
8	35815.3800000000000	Hamilton	Bermuda
9	35685.9800000000000	Bandar Seri Begawan	Brunei
10	34643.9600000000000	Schaan	Liechtenstein
11	34643.9600000000000	Vaduz	Liechtenstein
12	33236.8400000000000	George Town	Cayman Islands

To find GDP I divided GNP by population and then multiplied by 1,000,000 since the data is given in millions. Using WHERE I could specify that I only want to return results that I greater than (>) the average GDP. Then I could join the tables to return the GDP for each city.

20. **Display Columns with Limit (Rows 31-40):** *Scenario:* A market research firm requires detailed information on cities beyond the top rankings for a comprehensive analysis. You're tasked with

providing data on cities ranked between 31st and 40th by population to ensure a thorough understanding of urban demographics.

I used ORDER BY in descending order to order the table by population largest to smallest. Then I used OFFSET to skip over the first 30 rows and FETCH to display only the next 10 rows after that.

```
120
121   --20
122   SELECT *
123   FROM dbo.city as city
124   ORDER BY city.population desc
125   OFFSET 30 ROWS
126   FETCH NEXT 10 ROWS ONLY
```

Results Messages

	ID	Name	CountryCode	District	Population
1	1896	Shenyang …	CHN	Liaoning	4265200
2	1897	Kanton [Guangzhou] …	CHN	Guangdong	4256300
3	3208	Singapore …	SGP	-	4017733
4	3769	Ho Chi Minh City …	VNM	Ho Chi Minh City	3980000
5	1027	Chennai (Madras) …	IND	Tamil Nadu	3841396
6	2332	Pusan …	KOR	Pusan	3804522
7	3794	Los Angeles …	USA	California	3694820
8	150	Dhaka …	BGD	Dhaka	3612850
9	3068	Berlin …	DEU	Berliini	3386667
1…	2710	Rangoon (Yangon) …	MMR	Rangoon [Yangon]	3361700

QUERIES LIST

```
--1
SELECT COUNT(city.Name)
FROM dbo.city AS City
WHERE CountryCode LIKE 'USA'

--2
SELECT TOP 1(country.LifeExpectancy), *
FROM dbo.country AS country

--3
SELECT *
FROM dbo.city AS city
WHERE city.name LIKE '%new%'

--4
SELECT TOP 10 (city.population), *
FROM dbo.city AS city
ORDER BY city.population desc

--5
SELECT *
FROM dbo.city AS city
WHERE city.population > 2000000
ORDER BY Population desc

--6
SELECT *
```

```
FROM dbo.city as city
WHERE City.name LIKE 'be%'

--7
SELECT *
FROM dbo.city AS city
WHERE city.population <1000000 AND city.population >500000
ORDER BY city.population desc

--8
SELECT *
FROM dbo.city AS city
ORDER BY city.name

--9
SELECT TOP 1 (city.population), *
FROM dbo.city AS city

--10
SELECT COUNT(DISTINCT city.name) AS namecount , city.name
FROM dbo.city AS city
GROUP BY city.name
ORDER BY city.name

--11

SELECT TOP 1(city.population), *
FROM dbo.city AS city
ORDER BY city.population

--12

SELECT TOP 1 (country.population), *
FROM dbo.country AS country
ORDER BY country.population desc

--13

SELECT *
FROM dbo.city AS city
JOIN dbo.country AS country
ON city.countrycode = country.code
WHERE Country.Capital = city.ID AND city.countrycode LIKE 'ESP'

--14

SELECT TOP 1(country.LifeExpectancy), *
FROM dbo.country AS country
ORDER BY country.lifeexpectancy desc

--15
SELECT city.name
FROM dbo.City AS city
JOIN dbo.country AS country
ON city.countrycode = country.code
WHERE country.region LIKE '%europe%'
ORDER BY city.name

--16

SELECT AVG(city.population) as avgcitypopulation, country.name,
FROM dbo.city as city
JOIN dbo.country as country
```

```
ON city.countrycode = country.code
GROUP BY country.name

--17
SELECT city.name, city.population
FROM dbo.city
JOIN dbo.country as country
ON city.countrycode = country.code
WHERE country.capital = city.ID
ORDER BY city.population desc

--18

SELECT country.name, (country.population / country.Surfacearea) AS PopulationDensity
FROM dbo.country as country
ORDER BY (country.population / country.Surfacearea)

--19
SELECT round((country.gnp/country.population)*1000000,2) AS GDP, city.name AS CityName , country.name AS
CountryName
FROM dbo.country as country
JOIN dbo.city as city
ON city.countrycode = country.code
WHERE country.population > 0
and country.gnp > 0
and (country.gnp/country.population) > (SELECT AVG(country.gnp/country.population) FROM dbo.country where
population >0)
ORDER BY GDP desc



--20
SELECT *
FROM dbo.city as city
ORDER BY city.population desc
OFFSET 30 ROWS
FETCH NEXT 10 ROWS ONLY
```

Bibliography

Accelebrate (2022) *Azure SQL vs. SQL Server: How to choose the right database*, *Accelebrate*. https://www.accelebrate.com/blog/azure-sql-versus-sql-server.