

PROJECT: UNCOVERING THE WORLD'S OLDEST BUSINESSES



Staffelter Hof Winery is Germany's oldest business, established in 862 under the Carolingian dynasty. It has continued to serve customers through dramatic changes in Europe, such as the Holy Roman Empire, the Ottoman Empire, and both world wars. What characteristics enable a business to stand the test of time?

To help answer this question, BusinessFinancing.co.uk researched the oldest company still in business in **almost** every country and compiled the results into several CSV files. This dataset has been cleaned.

Having useful information in different files is a common problem. While it's better to keep different types of data separate for data storage, you'll want all the data in one place for analysis. You'll use joining and data manipulation to work with this data and better understand the world's oldest businesses.

The Data

businesses and new_businesses

Column	Description
business	Name of the business (varchar)
year_founded	Year the business was founded (int)
category_code	Code for the business category (varchar)
country_code	ISO 3166-1 three-letter country code (char)

countries

Column	Description
country_code	ISO 3166-1 three-letter country code (varchar)
country	Name of the country (varchar)
continent	Name of the continent the country exists in (varchar)

categories

Column	Description
category_code	Code for the business category (varchar)
category	Description of the business category (varchar)

 Projects Data DataFrame as oldest_business_continent

```
-- What is the oldest business on each continent?
WITH OBC AS (SELECT
  c.continent,
  c.country,
  b.business,
  ROW_NUMBER() OVER(PARTITION BY c.continent ORDER BY b.year_founded ASC ) AS rn,
  year_founded
FROM countries AS c
INNER JOIN businesses AS b
USING(country_code))

SELECT continent, country, business, year_founded
FROM OBC
WHERE rn=1;
```

index	...	↑↓	continent	...	↑↓	country	...	↑↓	business	...	↑↓	year_founded
			0			Africa			Mauritius			Mauritius Post
			1			Asia			Japan			Kongō Gumi
			2			Europe			Austria			St. Peter Stifts Kulinarium
			3			North America			Mexico			La Casa de Moneda de México
			4			Oceania			Australia			Australia Post
			5			South America			Peru			Casa Nacional de Moneda

Rows: 6

Expand

 Projects Data DataFrame as count

```
-- How many countries per continent lack data on the oldest businesses
-- Does including the `new_businesses` data change this?

SELECT
  continent,
  COUNT(*) AS countries_without_businesses
FROM countries AS c
WHERE c.country_code NOT IN (
  SELECT country_code
  FROM businesses
  UNION
  SELECT country_code
  FROM new_businesses
)
GROUP BY c.continent;
```

...	↑↓	contin...	...	↑↓	countries_without_businesses	...	↑↓
			0		Africa		3
			1		Asia		7
			2		South America		3
			3		Europe		2
			4		North America		5
			5		Oceania		10

Rows: 6

Expand

 Projects Data DataFrame as 0

```
-- Which business categories are best suited to last over the course of centuries?
WITH OBCC AS (
  SELECT
    co.continent AS continent,
    ca.category AS category,
    b.year_founded AS year_founded,
    ROW_NUMBER() OVER(PARTITION BY co.continent,ca.category ORDER BY b.year_founded ASC) AS rn
  FROM countries AS co
  INNER JOIN businesses AS b
  USING(country_code)
  INNER JOIN categories AS ca
  USING(category_code)
)
SELECT
  continent,
  category,
  year_founded
FROM OBCC
WHERE rn=1;
```

...	↑↓	contin...	...	↑↓	category	...	↑↓	year...	...	↑↓
0		Africa			Agriculture			1947		
1		Africa			Aviation & Transport			1854		
2		Africa			Banking & Finance			1892		
3		Africa			Distillers, Vintners, & Breweries			1933		
4		Africa			Energy			1968		
5		Africa			Food & Beverages			1878		
6		Africa			Manufacturing & Production			1820		
7		Africa			Media			1943		
8		Africa			Mining			1962		
9		Africa			Postal Service			1772		
10		Asia			Agriculture			1930		
11		Asia			Aviation & Transport			1858		
12		Asia			Banking & Finance			1830		
13		Asia			Cafés, Restaurants & Bars			1153		
14		Asia			Conglomerate			1841		
15		Asia			Construction			578		

Rows: 56

 Expand