Oldest Businesses Around the World – Longevity, Industry, and Historical Insights

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)

Data Validation and Initial Exploration

Projects Data DataFrame as `data_structure`

```
-- review the structure of all required tables.

SELECT
    table_name,
    column_name,
    data_type
FROM information_schema.columns
WHERE table_name IN ('businesses', 'countries', 'categories')
ORDER BY table_name, ordinal_position;
```

index	table_name	column_name	data_type
0	businesses	business	character varying
1	businesses	year_founded	integer
2	businesses	category_code	character varying
3	businesses	country_code	character
4	categories	category_code	character varying
5	categories	category	character varying
6	countries	country_code	character varying
7	countries	country	character varying
8	countries	continent	character varying

Rows: 9 Expand

Projects Data DataFrame as `b`

```
-- Overview of the table businesses

SELECT *
FROM businesses
LIMIT 5;
```

	business	year…	catego…	coun…
0	Hamoud Boualem	1878	CAT11	DZA
1	Communauté Électrique du Bénin	1968	CAT10	BEN
2	Botswana Meat Commission	1965	CAT1	BWA
3	Air Burkina	1967	CAT2	BFA
4	Brarudi	1955	CAT9	BDI

Rows: 5 Expand

Projects Data DataFrame as `n`

```
-- Overview of the table businesses

SELECT *
FROM new_businesses;
```

	business	year…	catego…	coun…
0	Fiji Times	1869	CAT13	FJI
1	J. Armando Bermúdez & Co.	1852	CAT9	DOM

Rows: 2 Expand

Projects Data DataFrame as

```sql
-- Overview of the table countries

SELECT *
FROM countries
LIMIT 5;
```

	coun...	country	c.
0	AFG	Afghanistan	Asia
1	AGO	Angola	Africa
2	ALB	Albania	Europe
3	AND	Andorra	Europe
4	ARE	United Arab Emirates	Asia

Rows: 5 Expand

Projects Data DataFrame as

```sql
-- Overview of the table categories

SELECT *
FROM categories
LIMIT 5;
```

	catego...	category
0	CAT1	Agriculture
1	CAT2	Aviation & Transport
2	CAT3	Banking & Finance
3	CAT4	Cafés, Restaurants & Bars
4	CAT5	Conglomerate

Rows: 5 Expand

Projects Data DataFrame as

```sql
-- Check referential integrity

SELECT DISTINCT category_code
FROM businesses
WHERE category_code NOT IN (
    SELECT category_code FROM categories
);
```

Your query ran successfully but returned no results.

Projects Data DataFrame as

```sql
-- Check country codes match

SELECT DISTINCT country_code
FROM businesses
WHERE country_code NOT IN (
    SELECT country_code FROM countries
);
```

Your query ran successfully but returned no results.

Projects Data DataFrame as

```
SELECT
    MIN(year_founded) AS min_year_founded,
    MAX(year_founded) AS max_year_founded
FROM businesses;
```

	min_year_f...	max_year_f...
0	578	1999

Rows: 1 Expand

Projects Data DataFrame as

```
SELECT
    business,
    year_founded,
    category_code,
    country_code
FROM businesses
WHERE year_founded IS NULL
    AND category_code IS NULL
    AND country_code IS NULL;
```

Your query ran successfully but returned no results.

Projects Data DataFrame as

```
SELECT
    country,
    continent
FROM countries
WHERE country IS NULL
    AND continent IS NULL;
```

Your query ran successfully but returned no results.

Projects Data DataFrame as

```
SELECT
    category
FROM categories
WHERE category IS NULL;
```

Your query ran successfully but returned no results.

Data Coverage Gaps: Countries Without Identified Oldest Businesses

Projects Data DataFrame as

```sql
-- Identify countries that do not have any recorded oldest business even after combining existing and newly added business data

SELECT
    c.continent,
    COUNT(c.country) AS countries_without_businesses
FROM countries AS c
LEFT JOIN (
    SELECT *
    FROM businesses
    UNION ALL
    SELECT *
    FROM new_businesses
    WHERE business IS NULL
) AS b
USING (country_code)
WHERE b.business IS NULL
GROUP BY c.continent
ORDER BY countries_without_businesses DESC;
```

	contin...	countries_without_businesses
0	Oceania	11
1	Asia	7
2	North America	6
3	Africa	3
4	South America	3
5	Europe	2

Rows: 6 Expand

Longevity of Business Categories Across Continents

Projects Data DataFrame as

```sql
-- Find the earliest founding year for each business category within each continent to assess long-term survivability

SELECT
    c1.continent,
    c2.category,
    MIN(b.year_founded) AS earliest_founding_year
FROM businesses AS b
JOIN countries AS c1
    ON b.country_code = c1.country_code
JOIN categories AS c2
    ON b.category_code = c2.category_code
GROUP BY
    c1.continent,
    c2.category
ORDER BY earliest_founding_year ASC;
```

	contin...	category	earliest_founding_year
0	Asia	Construction	578
1	Europe	Cafés, Restaurants & Bars	803
2	Europe	Distillers, Vintners, & Breweries	862
3	Europe	Manufacturing & Production	864
4	Asia	Cafés, Restaurants & Bars	1153
5	Europe	Agriculture	1218
6	Europe	Tourism & Hotels	1230
7	Europe	Mining	1248
8	Europe	Medical	1422
9	Europe	Postal Service	1520
10	North America	Manufacturing & Production	1534
11	South America	Banking & Finance	1565
12	Asia	Tourism & Hotels	1584
13	Europe	Banking & Finance	1606
14	South America	Manufacturing & Production	1621
15	North America	Agriculture	1638

Rows: 56 Expand

Global Distribution of Businesses Founded Before the Year 1000

Projects Data DataFrame as

```sql
-- Count businesses founded before the year 1000 grouped by continent

SELECT
    c.continent,
    COUNT(b.business) AS num_of_businesses
FROM businesses b
JOIN countries c
    ON b.country_code = c.country_code
WHERE b.year_founded < 1000
GROUP BY c.continent
ORDER BY num_of_businesses DESC;
```

	c.	num_of_busi...
0	Europe	5
1	Asia	1

Rows: 2 Expand

businesses_before_year_1000

Comparing the Average Age of Businesses Across Continents

Projects Data DataFrame as

```
-- Calculate the average age of businesses per continent [Business age = current year - year founded]

SELECT
    c.continent,
    ROUND(
        AVG(EXTRACT(YEAR FROM CURRENT_DATE) - b.year_founded)::NUMERIC,
        2
    ) AS avg_business_age
FROM businesses b
JOIN countries c
    ON b.country_code = c.country_code
GROUP BY c.continent
ORDER BY avg_business_age DESC;
```

	contin...	avg_busine...
0	Europe	516.43
1	South America	274.78
2	North America	198.47
3	Asia	187.12
4	Oceania	139
5	Africa	100.8

Rows: 6 Expand

avg_business_age

Geographic Reach of Business Categories Across Countries

Projects Data DataFrame as

```sql
-- Count the number of distinct countries in which each business category exists

SELECT
    cat.category_code,
    cat.category,
    COUNT(DISTINCT con.country) AS num_of_countries
FROM businesses b
JOIN countries con
    ON b.country_code = con.country_code
JOIN categories cat
    ON b.category_code = cat.category_code
GROUP BY cat.category_code, cat.category
ORDER BY num_of_countries DESC;
```

	catego…	category	num_of_co…
0	CAT3	Banking & Finance	37
1	CAT9	Distillers, Vintners, & Breweries	22
2	CAT2	Aviation & Transport	19
3	CAT16	Postal Service	16
4	CAT12	Manufacturing & Production	15
5	CAT13	Media	7
6	CAT1	Agriculture	6
7	CAT11	Food & Beverages	6
8	CAT4	Cafés, Restaurants & Bars	6
9	CAT17	Retail	4
10	CAT10	Energy	4
11	CAT19	Tourism & Hotels	4
12	CAT5	Conglomerate	3
13	CAT15	Mining	3
14	CAT7	Consumer Goods	3
15	CAT8	Defense	3

Rows: 19 Expand

Most Dominant Business Categories Within Each Continent

Projects Data DataFrame as

```sql
-- Rank business categories within each continent based on the number of businesses

SELECT
    con.continent,
    cat.category,
    COUNT(b.business) AS num_of_businesses,
    DENSE_RANK() OVER (
        PARTITION BY con.continent
        ORDER BY COUNT(b.business) DESC
    ) AS rank
FROM businesses b
JOIN countries con
    ON b.country_code = con.country_code
JOIN categories cat
    ON b.category_code = cat.category_code
GROUP BY con.continent, cat.category
ORDER BY con.continent, num_of_businesses DESC;
```

	contin...	category	num_of_busi...	
0	Africa	Banking & Finance	17	1
1	Africa	Aviation & Transport	10	2
2	Africa	Postal Service	9	3
3	Africa	Media	4	4
4	Africa	Distillers, Vintners, & Breweries	3	5
5	Africa	Agriculture	3	5
6	Africa	Manufacturing & Production	1	6
7	Africa	Food & Beverages	1	6
8	Africa	Mining	1	6
9	Africa	Energy	1	6
10	Asia	Aviation & Transport	7	1
11	Asia	Banking & Finance	6	2
12	Asia	Manufacturing & Production	3	3
13	Asia	Cafés, Restaurants & Bars	3	3
14	Asia	Retail	3	3
15	Asia	Conglomerate	3	3

Rows: 56 Expand

Oldest Operating Businesses by Continent (Tie-Aware Historical Analysis)

Projects Data DataFrame as

```sql
-- Rank businesses by founding year within each continent
-- Using RANK() to preserve ties in founding years

WITH ranked_businesses AS (
    SELECT
        con.continent,
        b.business,
        cat.category,
        b.year_founded,
        RANK() OVER (
            PARTITION BY con.continent
            ORDER BY b.year_founded ASC
        ) AS rank
    FROM businesses b
    JOIN countries con
        ON b.country_code = con.country_code
    JOIN categories cat
        ON b.category_code = cat.category_code
)
SELECT
    continent,
    business,
    category,
    year_founded AS oldest_year
FROM ranked_businesses
WHERE rank = 1
ORDER BY continent;
```

	contin…	business	category	old…
0	Africa	Mauritius Post	Postal Service	1772
1	Asia	Kongō Gumi	Construction	578
2	Europe	St. Peter Stifts Kulinarium	Cafés, Restaurants & Bars	803
3	North America	La Casa de Moneda de México	Manufacturing & Production	1534
4	Oceania	Australia Post	Postal Service	1809
5	South America	Casa Nacional de Moneda	Banking & Finance	1565

Rows: 6 Expand