**Forbes Companies Dataset (Excel Project)**

**Введение**

Данный проект содержит анализ топ 2000 компаний, взятых из датасета Forbes за 2022 год.

Проект создан с использованием программы Microsoft Excel с использованием:

* Сводных таблиц
* Визуализации на основе сводных таблиц
* Функции ЕСЛИМН

**Описание данных**

Датасет взят с сайта “Kaggle” и содержит ряд показателей:

* Позиция компании в рейтинге
* Название компании
* Отрасль деятельности компании
* Страна, в которой расположена штаб-квартира компании
* Год основания компании
* Исполнительный директор компании
* Выручка компании (млрд. $)
* Прибыль компании (млрд. $)
* Объем активов компании (млрд. $)
* Рыночная стоимость компании (млрд. $)
* Количество работников в компании
* Континент (добавлено с помощью функции ЕСЛИМН)

Формат датасета - CSV.

**Подготовка к анализу**

Изначально, тип данных в датасете некорректно распознавался Excel, потому что использовался знак точки для отделения дробной части. Данная проблема была решена с помощью опции «Найти и заменить». Знак точки был заменен на знак запятой. Далее, благодаря функции ЕСЛИМН был создан новый показатель – континент. В функцию было введено несколько условий, благодаря которым она присваивала определенный континент каждой компании, на основании локации штаб-квартиры (рис. 1).



Рис. 1 – Функция ЕСЛИМН

Затем, были построены сводные таблицы, а именно:

* Таблицы, связанные с отраслями: отрасли с наибольшим количеством компаний (лист Industry Quantity) и отрасли с наибольшим объемом средней выручки (лист Industry Revenue)
* Таблицы, связанные со странами: страны с наибольшим количеством компаний (лист Country Quantity) и страны с наибольшим объемом средней выручки (лист Country Revenue)
* Таблицы, связанные с континентами: распределение суммарного объема выручки (лист Continent Revenue), суммарной прибыли (лист Continent Profit), суммарных активов (лист Continent Assets), суммарной рыночной стоимости (лист Continent Market Value).

На основании всех сводных таблиц были построены соответствующие диаграммы: гистограммы для отраслей и стран, и круговые диаграммы для континентов (ввиду удобства отображения долей каждого континента). Также были созданы срезы стран и отраслей, благодаря которым можно посмотреть и проанализировать ситуацию по конкретным странам или отраслям. Срезы подключены ко всем сводным таблицам, поэтому при выборе конкретного значения в срезе, все визуализации изменяются автоматически.

Также, в дашборде имеется интерактивная карта мира, которая отражает суммарное количество работников в определенной отрасли. Значения по континентам отражены в виде пузырьков, которые также меняют свой размер, в зависимости от числа. Также, наибольшее значение из всех выделяется зеленым цветом.

**Заключение**

Исходя из результатов, можно сделать ряд выводов:

* Среди всех представленных компаний, представленных в датасете, наибольшее число компаний принадлежит банковской отрасли. Другие популярные отрасли – финансовая, строительная, сырьевая и нефтегазовая. Большинство компаний – американские, также много компаний зарегистрировано в Китае, Японии, Южной Корее и Великобритании.
* Наибольшую среднюю выручку приносят немецкие, французские, саудовские, британские и российские компании. Отрасли с наибольшей средней выручкой – розничная, нефтегазовая, пищевая, торговая и отрасль потребительских товаров длительного назначения.
* Азия и Северная Америка лидируют по суммарным объемам выручки, прибыли, активов и капитализации. Европа также имеет существенную долю. Доля Океании, Африки и Южной Америки незначительна.
* Азия и Северная Америка также лидируют по количеству работников, занятых в различных отраслях.

Данная информация может быть полезна при создании бизнес-плана: самые передовые компании отражают примерную ситуацию на рынке как отдельной страны, так и всех стран, представленных в данных. Поэтому, информация из данного проекта может быть полезна при выборе рынка для новой компании и отрасли для нее.