

Calculating the NPV of the instruments

Чиркова Юлия, Столярова Анастасия

November 2023

1 Цели проекта

Целью данного проекта является расчет чистой приведенной стоимости финансовых инструментов, поиск ошибок среди предложенных данных для инструментов и поиск комплекса мер, разрешающих возникающую проблему.

2 Используемые теоретические концепции и методы

NPV (Net Present Value) или чистая приведенная стоимость – это финансовый показатель, который позволяет оценить стоимость инвестиционного проекта или финансового инструмента. Он вычисляется путем приведения будущих денежных потоков к их текущей стоимости с использованием дисконтирования.

Для расчета NPV необходимо выполнить следующие шаги:

1. Определить предполагаемые денежные потоки, которые будут получены от инвестиционного проекта. Обычно это включает начальные инвестиции, прогнозируемые доходы или расходы на протяжении срока проекта и оценку стоимости реализации проекта в конце срока.
2. Определить ставку дисконтирования, которая отражает желаемую доходность инвестора от проекта. Эта ставка должна учитывать риски проекта и альтернативные возможности вложения денег.
3. Привести каждый денежный поток к текущему периоду и просуммировать дисконтированные денежные потоки, то есть использовать формулу:

$$NPV = \sum \frac{CF_t}{(1 + r)^t}$$

Где:

- NPV - чистая приведенная стоимость
- CF_t - денежный поток в момент времени t
- r - ставка дисконта
- t - период времени

Полученное значение будет являться NPV.

Если NPV положительный, это означает, что инвестиционный проект приносит положительную экономическую прибыль и может быть выгодным. Если NPV отрицательный, проект не оправдывает затрат и, вероятно, не является выгодным.

NPV также позволяет сравнить различные проекты: следует выбирать тот, у которого значение NPV наибольшее.

3 Используемые инструменты

В рамках данного проекта мы использовали библиотеки языка Python, позволяющие работать с большим объемом данных.

4 Задачи проекта

Реализацию проекта мы разбили на несколько этапов:

1. Предобработка данных
2. Поиск ошибок в данных
3. Поиск решения проблемы с данными
4. Расчет NPV(net present value) для каждого финансового инструмента
5. Сортировка полученных результатов по мере убывания их profitability и визуализация данных

4.1 Предобработка данных

Этап предобработки данных включал :

- 1 Выбор необходимых библиотек
- 2 Выгрузка данных

4.1.1 Выбор необходимых данных

Для работы в Google Collab с предлагаемым датасетом были подключены следующие библиотеки:

- *pandas* - для работы с датасетом
- *numpy* - для удобных математических функций для работы с векторами данных
- *seaborn* - для эстетичной визуализации результатов
- *matplotlib.pyplot* - для построения графиков

4.1.2 Выгрузка данных

Для работы с датасетом в среде Google Collab была использована функция из библиотеки *pandas*

```
data=pd.read_excel("Темы исследований.xlsx", sheet_name=6, skiprows=4)
```

Следующие строчки кода представляют данные в таблице в более удобном для работы виде: удаляют неинформативные строки и столбцы из таблицы, заполняют нулями нечисловые данные, задают названия столбцам.

4.2 Поиск ошибок в данных

В данном разделе мы хотим выяснить, есть ли в данных выбросы. Для этого изобразим распределение данных на диаграмме типа boxplot, чтобы понять, в каком диапазоне находятся наши денежные потоки, и отследить наличие значительно отличающихся сумм.

Отрисуем два графика: один для выплат агентов (тех сумм, которые они заплатили за покупку рассматриваемых финансовых инструментов), второй для денежных потоков, которые агенты получают впоследствии.

Гипотеза состоит в том, что первая группа - отрицательные денежные потоки, а вторая - положительные (поэтому удобнее анализировать их на двух различных графиках, а не на одном).

Для визуализации мы будем использовать график типа `boxplot` (также известный как "ящик с усами"). Напомним, как нужно читать такой график:

1. верхняя (правая) грань "ящика" показывает верхний квартиль (то значение, ниже которого находятся 75% выборки);
2. нижняя (левая) грань "ящика" показывает нижний квартиль (то значение, выше которого находятся 75% выборки);
3. черта внутри "ящика" это медиана, т.е. 50% выборки выше данного значения и 50% выборки - ниже;
4. верхняя (правая) засечка показывает максимальное значение, а все, что выше нее - выбросы;
5. нижняя (левая) засечка показывает минимальное значение, а все, что ниже нее - выбросы.

Предварительно нам потребуется сделать два массива данных для построения 2 графиков.

Алгоритм построения графиков для анализа выбросов:

1. Создается массив `payments` с помощью функции `np.array()`. В этом массиве сохраняются значения из первой строки исходного датасета `data`, начиная со второй колонки.
2. Создается массив `incomes` с помощью функции `np.array()`. В этом массиве сохраняются значения из исходного датасета `data` начиная со второй строки и второй колонки, после чего он преобразуется в одномерный массив с помощью функции `flatten()`.
3. Создается фигура с помощью функции `plt.figure()`, задаются размеры фигуры с помощью аргумента `figsize`.
4. Создается макет с помощью функции `plt.subplots.adjust()`, задается вертикальное расстояние между подграфиками с помощью аргумента `hspace`.
5. Создается первый подграфик с помощью функции `plt.subplot()`. В нем создается ящик с усами с помощью функции `sns.boxplot()`, где по оси `x` отображаются значения массива `payments`. Далее задаются заголовок подграфика, название оси `x`, деления оси `x` и сетка.
6. Создается второй подграфик аналогичным образом, но уже используется массив `incomes` и отображаются соответствующие значения.
7. С помощью функции `plt.show()` выводится фигура с подграфиками.

Заметим, что распределение плат за инструменты находится в диапазоне от $[-140, -105]$, при этом большая часть значений находится в еще более узком диапазоне.

Однако, в денежных потоках, получаемых агентом, есть значительное значение. Большинство выплат концентрируются около нуля, с учетом текущего масштаба графика, но есть выброс около 10000. Для нахождения номера столбца и строки с этим значением в нашем датасете можно использовать функцию `where()`.

4.3 Поиск решения проблемы с данными

Один из наиболее распространенных методов обработки выбросов - это замена выброса на медианное значение. Медиана - это значение, стоящее в середине некоторого отсортированного списка. Медиана является более репрезентативной характеристикой выборки, чем среднее значение, когда в данных присутствуют выдающиеся значения.

Приведенный код создает новый объект «`DataFrame`» с именем `med`, используя данные из массива «`incomes`» и рассчитывает медиану значений в этом «`DataFrame`».

4.4 Расчет NPV(net present value) для каждого финансового инструмента

Для расчета чистой приведенной стоимости была взята ставка процента, равная 5%. Предварительно были рассчитаны и записаны в массив коэффициенты, на которые впоследствии будут делиться денежные потоки каждого периода.

Алгоритм расчета NPV для инструмента:

1. Все потоки приводятся к текущему периоду делением значения денежного потока в конкретном периоде на соответствующий коэффициент из рассчитанного массива хранения массива.
2. Все приведенные денежные потоки суммируются.
3. Результат (чистая приведенная стоимость конкретного финансового инструмента) записывается в новую таблицу.

Затем на основе измененных данных была рассчитана чистая приведенная стоимость финансовых инструментов.

Вычисление NPV оформлено в отдельную функцию. Она принимает два аргумента:

- «rate» - коэффициент дисконтирования, который используется для приведения будущих денежных потоков к текущей стоимости.
- «data» - таблица данных, где каждый столбец представляет денежные потоки для отдельного инструмента.

Внутри функции, для каждого инструмента, создается вспомогательный столбец «coeffs», содержащий коэффициенты дисконтирования для каждого года. Затем в цикле, происходит деление каждого столбца с денежными потоками на соответствующие коэффициенты, что приводит денежные потоки к текущей стоимости. Затем происходит суммирование приведенных денежных потоков и получение NPV для каждого инструмента. Названия инструментов и их соответствующие NPV сохраняются в массивы «names» и «profits» соответственно.

После цикла создается и сортируется результирующая таблица «result», в которой инструменты упорядочены по убыванию NPV.

Наконец, функция возвращает эту таблицу как результат своего выполнения.

4.5 Сортировка полученных результатов по мере убывания их profitability и визуализация данных

При помощи функции

```
result=result.sort_values(by="Profitability", ascending=False)
```

библиотеки Pandas рассчитанные чистые приведенные стоимости финансовых инструментов сортируются в порядке убывания.

Для визуализации данных создается столбчатая диаграмма, отображающая прибыльность различных финансовых инструментов. Устанавливается заголовок графика, подписи осей, метки на оси X, и сетка при помощи функций библиотеки Seaborn.

5 Результаты

В результате работы описанного кода получили чистые приведенные стоимости для каждого финансового инструмента из предложенного датасета. Ознакомиться с численными результатами и непосредственно с кодом, решающим поставленную задачу, можно в проекте в Google Collaboratory.