



Обновление с 6

Администратор создал6 месяцев назад

35efa149

README.md 4,31 КБ

День 00. Отправляйтесь в учебный лагерь

Статистика под рукой

Содержание

1. [Глава I](#)
- 1.1. [Основные правила](#)
2. [Глава II](#)
- 2.1. [Правила дня](#)
3. [Глава III](#)
- 3.1. [вступление](#)
4. [Глава IV](#)
- 4.1. [Упражнение 00: Анскомб](#)

Глава I

Основные правила

- Ваши программы не должны завершаться неожиданно (выдавая ошибку при правильном вводе). Если это произойдет, ваш проект будет считаться неработоспособным и получит 0 баллов при оценке.
- Мы рекомендуем вам создать тестовые программы для вашего проекта, даже если эту работу не нужно будет отправлять и она не будет оцениваться. Это даст вам возможность легко проверить свою работу и работу ваших коллег. Вы найдете эти тесты особенно полезными во время вашей защиты. Действительно, во время защиты вы можете использовать свои тесты и/или тесты коллеги, которого вы оцениваете.
- Отправьте свою работу в назначенный репозиторий git. Оцениваться будет только работа в репозитории git.
- Если ваш код использует внешние зависимости, он должен использовать [модули Go](#) для управления ими.

Глава II

Правила дня

- Вы должны только сдать *.go файлы и (в случае внешних зависимостей) go.mod + go.sum
- Ваша программа должна принимать на свой стандартный ввод последовательность чисел, разделенных символами новой строки. Одно число также является последовательностью.
- Вы можете предположить, что он должен работать только с целыми числами (хотя вывод может быть с плавающей точкой, округленным до 2 знаков после запятой).
- Тем не менее, он должен выводить значимое сообщение об ошибке без паники во время выполнения, если подается какой-то неожиданный ввод, скажем, число вне границ, буквенные символы или пустая строка.
- Ваш код для этой задачи должен быть собран с помощью всего лишь go build

Глава III

вступление

Go обычно не считается языком науки о данных. Но это не значит, что он не может обрабатывать цифры. Фактически, он сравним с C в основных задачах. Кроме того, его может быть намного проще писать, отчасти потому, что GC управляет памятью, а стандартная библиотека Go довольно хороша. Нас постоянно учат, что доверять чутью при работе с важными данными — плохая идея. Чтобы сделать выборку чисел орлом или решкой, обычно лучше использовать статистический подход. Данные иногда тоже могут быть обманчивыми, как, например, [квартет Анскомба](#) , но чем больше метрик мы получим — тем более взвешенное решение мы сможем принять в итоге, не так ли?

Глава IV

Упражнение 00: Анскомб

So, let's say we have a bunch of integer numbers, strictly between -100000 and 100000. It may probably be a large set, so let's assume our application will read it from a standard input, separated by newlines. Right now let's think of four major statistical metrics that we can derive from this data, so by default we can print all of them as a result, for example, like this:

```
Mean: 8.2
Median: 9.0
Mode: 3
SD: 4.35
```

The order and actual format doesn't really matter as long as we can understand which is which. A couple of notes, though:

- 1. Input data may or may not be sorted. You don't need to write your own sorting algorithm, luckily, Go already has one in standard library and it works for integers.
- 2. Median is a middle number of a sorted sequence if its size is odd, and an average between two middle ones if their count is even.
- 3. Мода — это число, которое встречается чаще всего, и если их несколько, возвращается наименьшее из них. Вы можете подумать об использовании некоторой структуры для хранения подсчета чисел, и может помочь какой-нибудь стандартный контейнер Go.
- 4. Вы можете использовать как совокупность, так и обычное стандартное отклонение, в зависимости от того, что вы предпочитаете.
- 5. Называть кого-то «средним» может быть грубо.

Для пользователя также имеет смысл иметь возможность выбирать, какой из этих четырех параметров печатать, поэтому необходимо реализовать и это. По умолчанию это все, но должен быть способ указать, печатать ли только одну, две или три конкретные метрики из четырех при запуске программы (без перекомпиляции). В стандартной библиотеке есть встроенный модуль, позволяющий парсить дополнительные параметры.