Добрый день! Помогите разобраться :eyes::flashlight:.  
Карплю над 27 вопросом.  
Вроде все верно сделал, но ответ неправильный.Сделал предварительно  str(s).split('|') и  .explode("cast"). Во всех предыдущих заданиях работало правильно.*# эксплодим актеров*  
cast\_expl = data.copy()  
cast\_expl["cast"] = cast\_expl["cast"].apply(lambda s: str(s).split('|'))  
cast\_expl = cast\_expl.explode("cast")*# Создаем итерируемый объект combinations из ячеек нужного столбца,*  
*# помещаем его в словарь-счтетик Counter и находим самые частые комбинации.*  
Counter(combinations(cast\_expl.cast, 2)).most\_common(1)Ответ есть, но неправильный :cry::  
[(('Mark Wahlberg', 'Samuel L. Jackson'), 411)] (edited)

8 replies



[**Sergey Meshcheryakov**](https://app.slack.com/team/U0192UB6HMF)  [21 days ago](https://skillfactorydst.slack.com/archives/C01843G3S22/p1629036114049000?thread_ts=1629032970.048300&cid=C01843G3S22)

Добрый день.  
После cast\_expl = cast\_expl.explode("cast") в столбце 9438 имен (многие из которых повторяются). Из этого набора имен Вы составляете пары. При этом пары не имеют связи с фильмами (большинство из них) - это просто все возможные комбинации элементов данного массива.



[**Sergey Meshcheryakov**](https://app.slack.com/team/U0192UB6HMF)  [21 days ago](https://skillfactorydst.slack.com/archives/C01843G3S22/p1629037531049200?thread_ts=1629032970.048300&cid=C01843G3S22)

Вам нужно составить пары из актерского состава каждого фильма.

cast\_expl = data.copy()

cast\_expl["cast"] = cast\_expl["cast"].apply(lambda s: str(s).split('|'))

cast\_expl["cast"] = cast\_expl["cast"].apply(lambda s: list(combinations(sorted(s),2)))

После чего, Вам нужно посчитать, сколько раз, та или иная пара присутствует в датасете (т.е. получим количество фильмов, в которых снималась отдельно взятая пара). Можно это сделать 2 способами:

* explode + value\_counts
* конкатенация значений всех ячеек столбца cast + передача результата в Counter



[**Andrei Kukunov**](https://app.slack.com/team/U024DUN2HM3)  [21 days ago](https://skillfactorydst.slack.com/archives/C01843G3S22/p1629037885049400?thread_ts=1629032970.048300&cid=C01843G3S22)

Да, увидел, пары не имеют связи с фильмами.  Делаю вот так.  
Получаются комбинации по каждой ячейке:  
cast\_pairs = data["cast"].apply(lambda s: str(s).split('|')).apply(lambda i: list(combinations(i, 2)))Но далее вставляю его в Counter -- и не принимает: unhashable type: 'list'  
Counter(cast\_pairs).most\_common(3)пытаюсь вот так -- тоже:  
Counter(**hash**(cast\_pairs)).most\_common(3) (edited)



[**Sergey Meshcheryakov**](https://app.slack.com/team/U0192UB6HMF)  [21 days ago](https://skillfactorydst.slack.com/archives/C01843G3S22/p1629038059049700?thread_ts=1629032970.048300&cid=C01843G3S22)

***конкатенация****значений всех ячеек столбца cast*

Counter(cast\_expl["cast"].values.sum()).most\_common(5)

+ Вы потеряли сортировку перед получением комбинаций

(lambda s: list(combinations(sorted(s),2)))

(edited)



[**Andrei Kukunov**](https://app.slack.com/team/U024DUN2HM3)  [21 days ago](https://skillfactorydst.slack.com/archives/C01843G3S22/p1629038113050100?thread_ts=1629032970.048300&cid=C01843G3S22)

спасибо, только что увидел все ваши ответы сейчас изучу... (edited)



[**Sergey Meshcheryakov**](https://app.slack.com/team/U0192UB6HMF)  [21 days ago](https://skillfactorydst.slack.com/archives/C01843G3S22/p1629039905050500?thread_ts=1629032970.048300&cid=C01843G3S22)

Если что-то не получится, можно почитать обсуждения в треде этого сообщения  
<https://skillfactorydst.slack.com/archives/C01843G3S22/p1616245949022100?thread_ts=1616245949.022100&cid=C01843G3S22>

Андрей Соколов

\*Вопрос 27\*. Какие актеры чаще всего снимаются в одном фильме вместе  
Долго работает. Я делал вот так:  
all\_actors = []for line in data['cast']:  
    actors = line.split('|')  
    for i in actors:  
Show more

[Thread in #0\_project\_1-миллионером](https://skillfactorydst.slack.com/archives/C01843G3S22/p1616245949022100?thread_ts=1616245949.022100&amp;cid=C01843G3S22) | [Mar 20th](https://skillfactorydst.slack.com/archives/C01843G3S22/p1616245949022100?thread_ts=1616245949.022100&amp;cid=C01843G3S22) | [View message](https://skillfactorydst.slack.com/archives/C01843G3S22/p1616245949022100?thread_ts=1616245949.022100&amp;cid=C01843G3S22)



[**Andrei Kukunov**](https://app.slack.com/team/U024DUN2HM3)  [21 days ago](https://skillfactorydst.slack.com/archives/C01843G3S22/p1629041030050900?thread_ts=1629032970.048300&cid=C01843G3S22)

Получается, самые частые -- 3 пары с частотностью по 8 раз. Первую пару я видел и в других ветках. Она совпадает с одним из вариантов ответов. Так и  должно быть?  
Код такой:  
cast\_pairs = data["cast"].apply(lambda s: str(s).split('|')).apply(lambda i: list(combinations(**sorted**(i), 2))).values.sum()  
Counter(cast\_pairs).most\_common(5) (edited)

:raised_hands:1



[**Sergey Meshcheryakov**](https://app.slack.com/team/U0192UB6HMF)  [21 days ago](https://skillfactorydst.slack.com/archives/C01843G3S22/p1629041436051300?thread_ts=1629032970.048300&cid=C01843G3S22)

Да, указываете ту пару, что совпала с вариантом ответа