Навигация

[Описание данных 2](#_Toc27942669)

[1. World Happiness. 2](#_Toc27942670)

[2. Kickstarter projects 3](#_Toc27942671)

[Разведывательный анализ 4](#_Toc27942672)

[1. World Happiness 4](#_Toc27942673)

[2. Kickstarter projects 8](#_Toc27942674)

# Описание данных

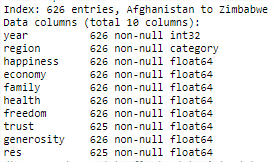
Анализ проводится на 2 разных датасетах: World Happiness и Kickstarter projects.

## World Happiness.

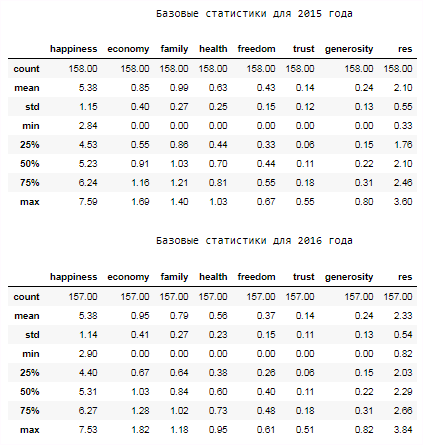
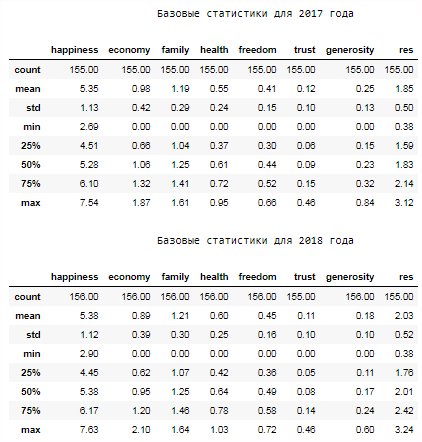
Данные с ежегодного исследования уровня счастья во всех странах, проводимого ООН. Данные взяты за 2015-2018 г. Датасет со странами, регионами и их уровнем счастья (*happiness*). Также указаны факторы, объясняющие уровень счастья: *economy* – реальное ВВП на душу населения, *family* – социальная поддержка, *health* - ожидаемая продолжительность здоровой жизни, *freedom* – свобода выбора в жизни, *trust* – восприятие коррупции (доверие государству), *generosity* – щедрость.

Ссылки на данные: [за 2015-2017г](https://www.kaggle.com/unsdsn/world-happiness), [за 2018г](https://www.kaggle.com/njlow1202/world-happiness-report-data-2018).

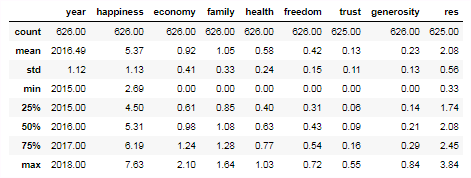
За 4 анализируемых года получаем следующий датасет:



В разбиении по отдельным годам базовые статистики выглядят следующим образом:

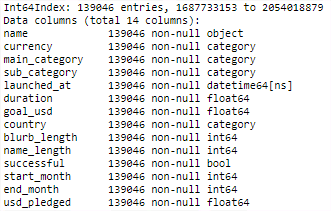
Базовые статистики суммарно за 4 года:

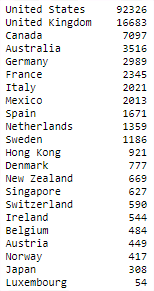


## Kickstarter projects

Датасет с информацией о проектах на Kickstarter за 2014-2019г. В анализе используется часть за 2015-2018. Имеет признаки: *id* – идентификатор проекта, name – название проекта, *currency* – валюта, *main\_category* – главная категория проекта, *sub\_category* – подкатегория проекта, *launched\_at* – время запуска кампании, *deadline* – время окончания кампании, *duration* – длительность кампании в днях, *goal\_usd* – цель в сборах (в долларах), *city* – город, *state* – штат, *country* – страна, *blurb\_length* – количество слов в описании кампании, *name\_length* – количество слов в названии проекта, *status* – статус проекта, *start\_month* – месяц запуска кампании, *end\_month* – месяц окончании кампании, *start\_Q* – квартал запуска, *end\_Q* – квартал окончания, *usd\_pledged* – сумма собранных средств к концу кампании.

Ссылка на данные: [кистартер](https://www.kaggle.com/yashkantharia/kickstarter-campaigns).

За анализируемый период представлено 139тыс. проектов.

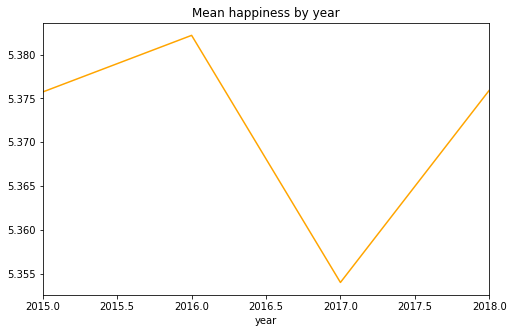


Основная часть проектов запущена в США, что объясняется спецификой самой платформы Kickstarter.

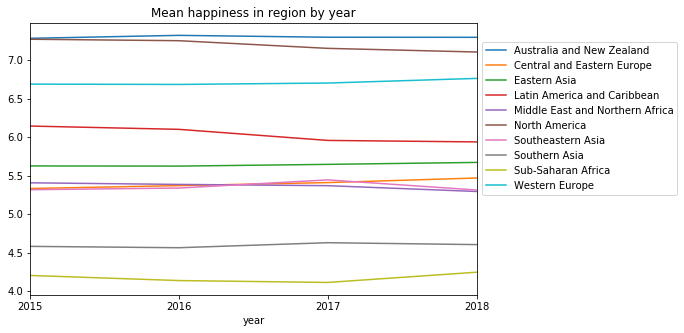
# Разведывательный анализ

## World Happiness

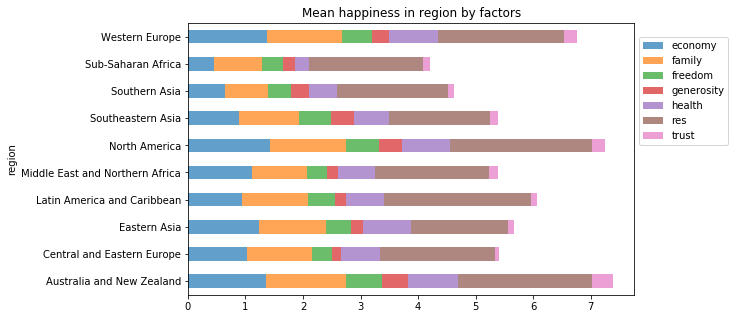
Средний уровень счастья для всех стран, динамика по годам: заметно снижение в 2017г. на 0.025



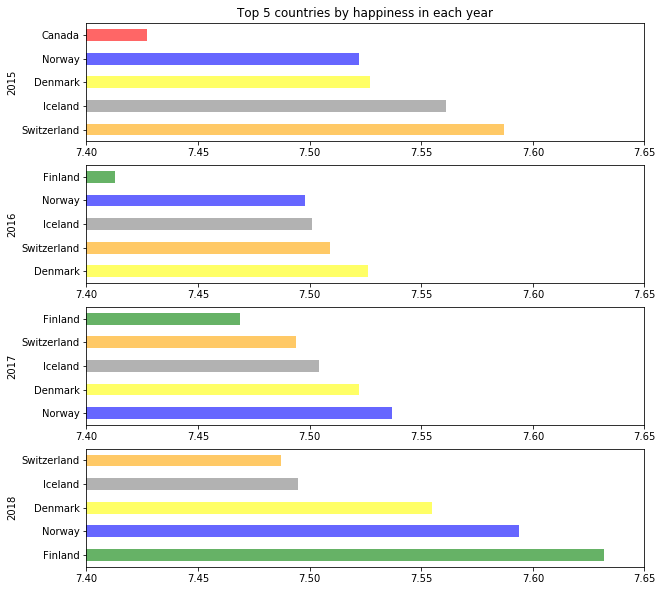
Средний уровень счастья для разных регионов по годам: в среднем страны Австралии и Новой Зеландии, а также Северной Америки более счастливы остальных регионов. Средние значения распределены в интервале от 4 до 7.5 баллов. Наименее счастливы страны Чёрной Африки – Африки к Югу от Сахары, и Южной Азии – к которой относятся страны Афганистан, Бангладеш, Бутан, Индия, Непал, Мальдивы, Пакистан, Шри-Ланка.



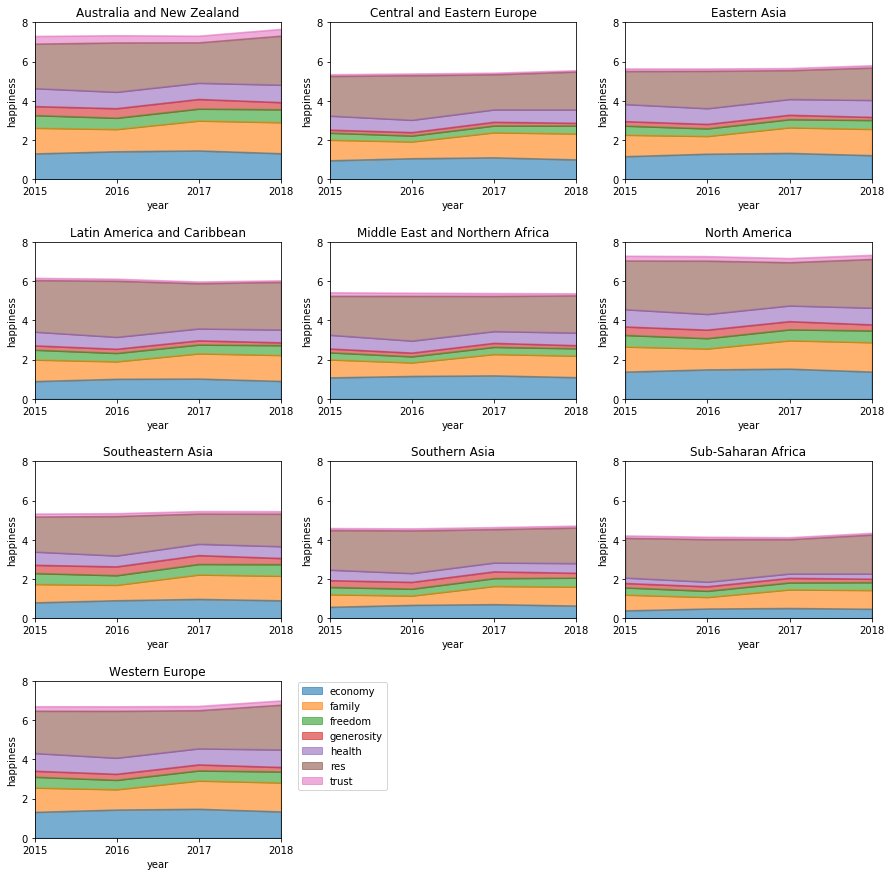
Средний уровень счастья для разных регионов по факторам: в странах Западной Европы, Северной Америки и Австралии и Новой Зеландии экономическое положение в среднем приносит больше счастья, чем в странах других регионов. В странах Африки к Югу от Сахары сравнительно мало счастья приносит такой фактор, как здоровье.



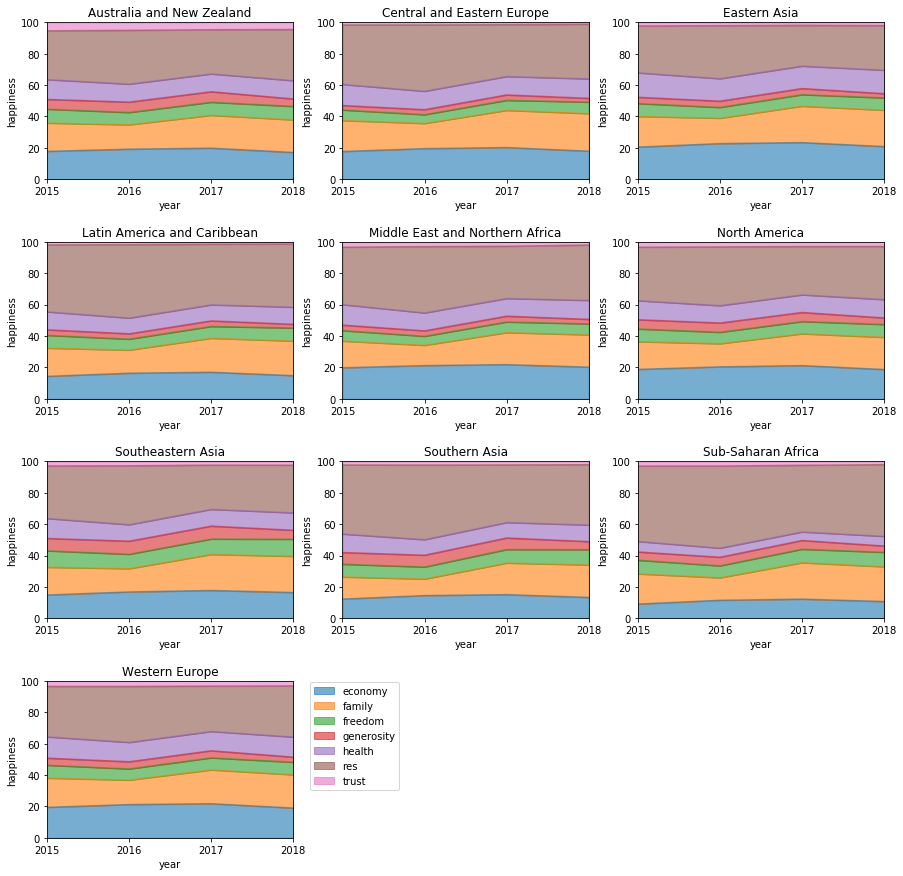
Топ-5 самых счастливых стран за каждый год: среди наиболее счастливых стран за все 4 года можно видеть Норвегию, Исландию, Данию. В целом состав не меняется, однако в 2015г Канада покинула топ-5 и её место заняла Финляндия, которая вошла в топ-5 в 2016г. и достигла 1 места к 2018г. Швейцария, обратно, спустилась с 1 места в 2015 на 5 место в 2018г.



Зависимость уровня счастья от разных факторов в среднем по каждому региону в динамике по годам: заметно, что Австралия и Новая Зеландия – в целом самый счастливый регион и задаёт максимальную планку в 7.5 баллов. Также заметно повышение счастья от семьи и социальной поддержки в 2017г почти по всем регионам. Остальные факторы не показывают значительных изменений.

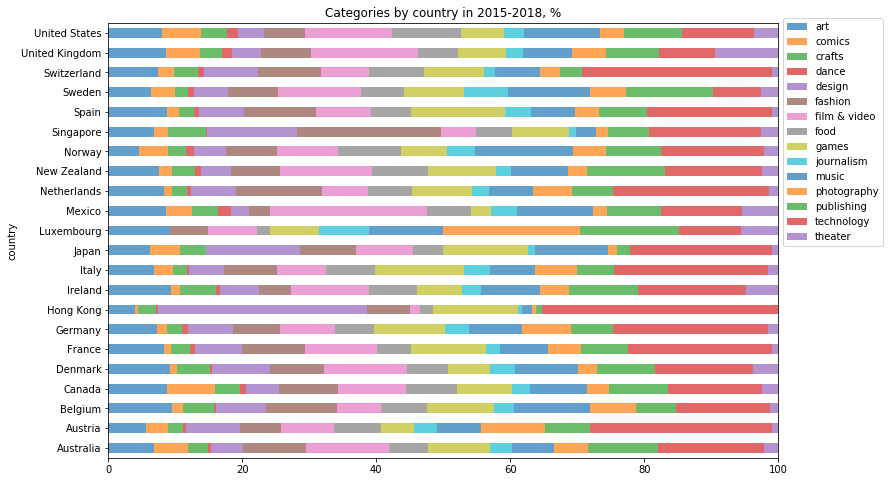


Доля разных факторов в общем уровне счастья в среднем по каждому региону в динамике по годам: заметно, что в исследовании Африки к Югу от Сахары, большая доля оценки счастья не приписана к какому-то определённому фактору и формирует собой необъяснённый остаток. В противоположность к этому – регион Восточной Азии, в котором доля необъяснённой оценки самая низкая.

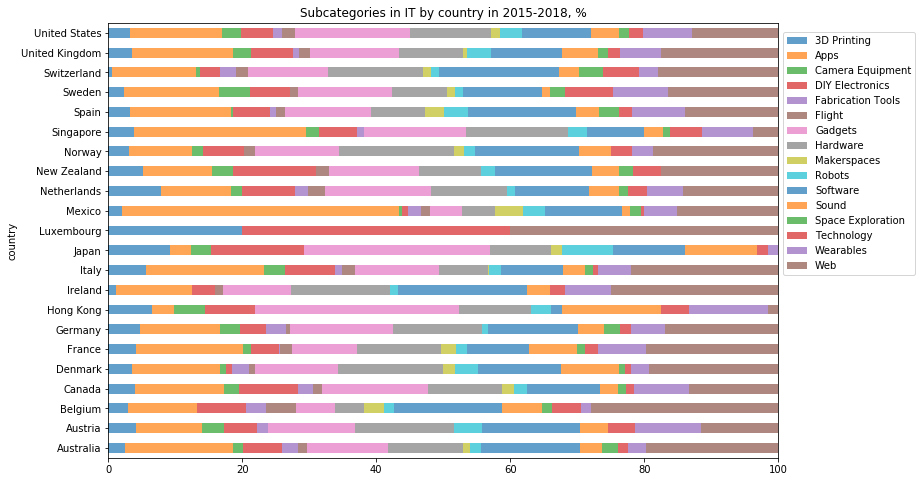


## Kickstarter projects

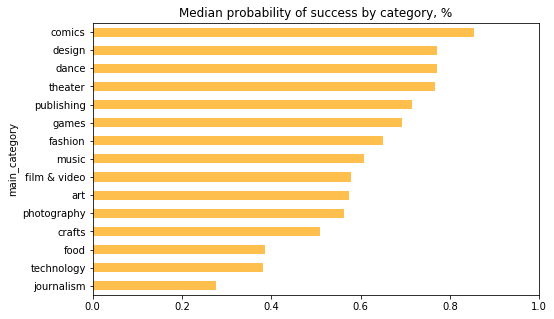
Доля проектов в каждой категории для каждой страны суммарно за 2015-2018г. В Канаде относительно большая доля проектов связана с комиксами, в Гонконге – с дизайном, в Сингапуре – с индустрией моды, в Мексике – с кино и видео индустриями, в Испании и Японии – с играми, в Люксембурге – с журналистикой и фотографией, в Швеции – с публицистикой, в Англии – с театром. ИТ проектов много в Швейцарии, Нидерландах, Гонконге и Австрии.



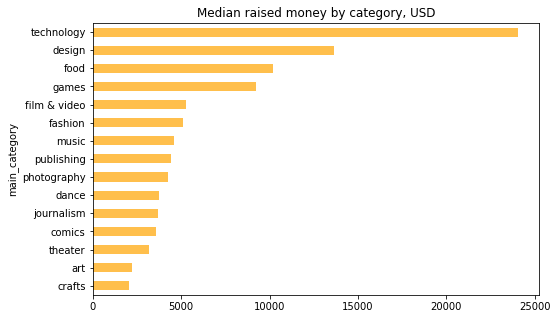
Доля проектов в ИТ в каждой субкатегории для каждой страны суммарно за 2015-2018г. В Люксембурге мало проектов в целом – 54 за все 4 года, однако в сфере ИТ они распределились в основном на 3д принтинг, сделай-сам электронику и веб-технологии. В Мексике большая доля проектов, связанных с приложениями, в Японии и Гонконге – с гаджетами, роботами и звуком, в Норвегии – с «железом», в Ирландии – с программным обеспечением, в Швеции – с технологиями, в Гонконге – с носимой электроникой, в Бельгии и Ирландии – с веб-технологиями и разработкой.



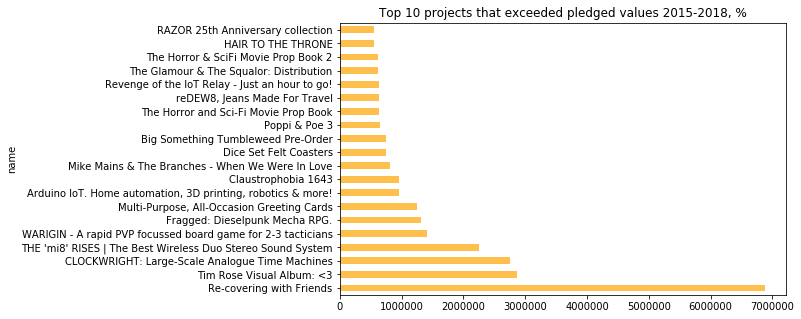
Медианная вероятность успеха кампании в разных категориях суммарно за 2015-2018г: самая высокая вероятность успеха у проектов в категории комиксы, самая низкая – в журналистике, технологиях и еде.



Медианная сумма вырученных средств в разных категориях суммарно за 2015-2018г: больше всего средств за один проект можно собрать в категории технологий, меньше всего – в категориях ремесла, искусства, театра, комиксов, журналистики и хореографии.



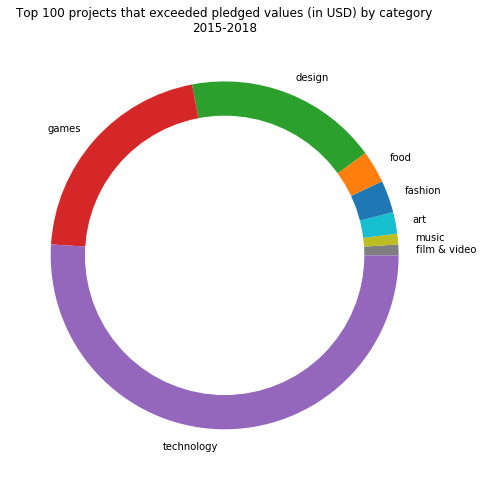
Топ-5 проектов, превысивших цели по сбору средств за 2015-2018г в %: проект «[Re-covering with Friends](https://www.kickstarter.com/projects/1343123793/re-covering-with-friends)» собрал почти в 70тыс раз больше средств, чем запланировал. Его цель была $1, а собрал он $68,764. Почти в 30тыс раз больше средств собрал проект «[Tim Rose Visual Album: <3](https://www.kickstarter.com/projects/timrosemusic/tim-rose-visual-album-3)» – $28,688 вместо $1. С похожим результатом – проект «[CLOCKWRIGHT: Large-Scale Analogue Time Machines](https://www.kickstarter.com/projects/clockwright/clockwright-large-scale-analogue-time-machines)», собравший $27,588 вместо $1.



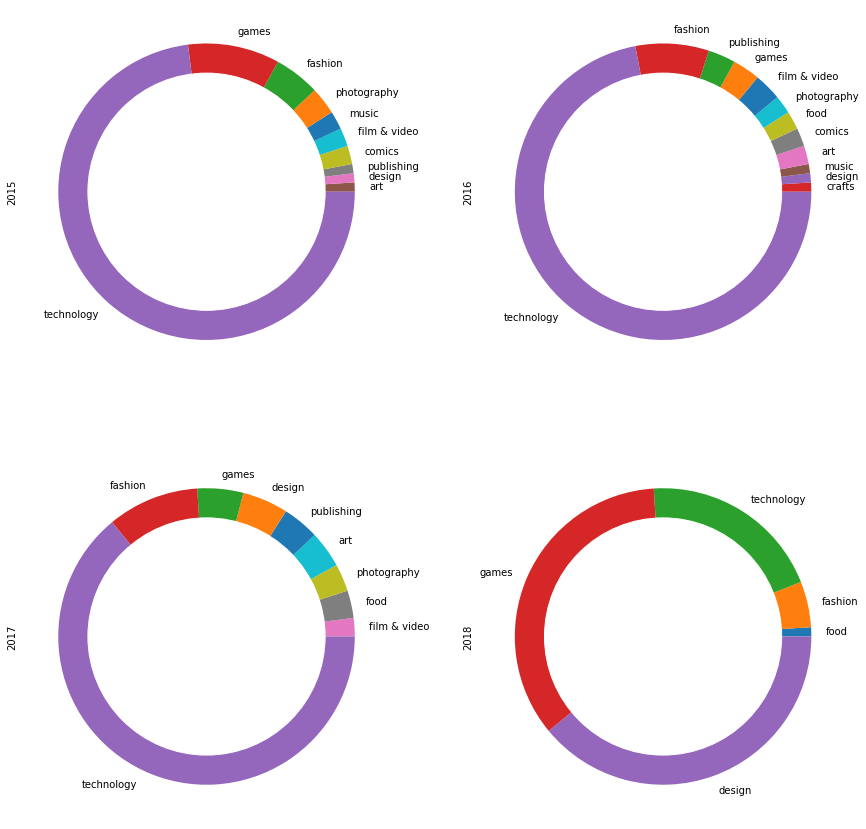
Топ-5 проектов, превысивших цели по сбору средств за 2015-2018г в долларах­­: игра «[Tainted Grail: The Fall of Avalon](https://www.kickstarter.com/projects/awakenrealms/tainted-grail-the-fall-of-avalon)» собрала $6,424,257 вместо $52,017. Проект «[ZeTime: World's first smartwatch with hands over touchscreen](https://www.kickstarter.com/projects/1282890542/zetime-worlds-first-smartwatch-with-hands-over-tou)» собрал $5,333,792 вместо $50,000.



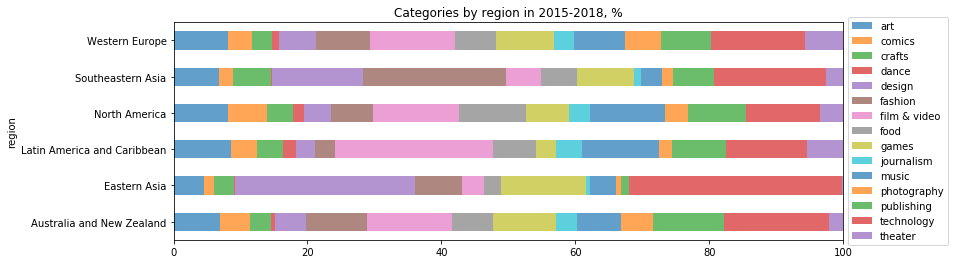
Доля категорий в топ-100 проектах, превысивших цели по сбору средств за 2015-2018г в долларах: значительно преобладают технологии, исчисляя половину самых прибыльных проектов. После них – игры и дизайн.



Доля категорий в топ-100 проектах, превысивших цели по сбору средств за каждый год 2015-2018 в долларах: в 2015-2016 году почти 70% проектов были из категории технологий, в 2017 эта доля снизилась до 60%, а в 2018 картина и вовсе изменилась – 75% проектов относятся в равной степени либо с дизайну, либо к играм, а технологии занимают остальные 15%.



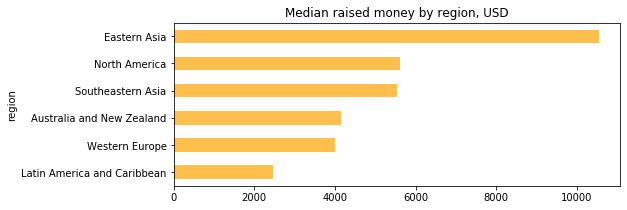
Доля проектов в каждой категории для каждого региона суммарно за 2015-2018г: в разбиении по регионам заметно, что Восточной Азии распространены проекты в категории дизайна и технологий, в Латинской Америке и Карибских островах – в категории фильмов и видео, в Юго-Восточной Азии – моды.



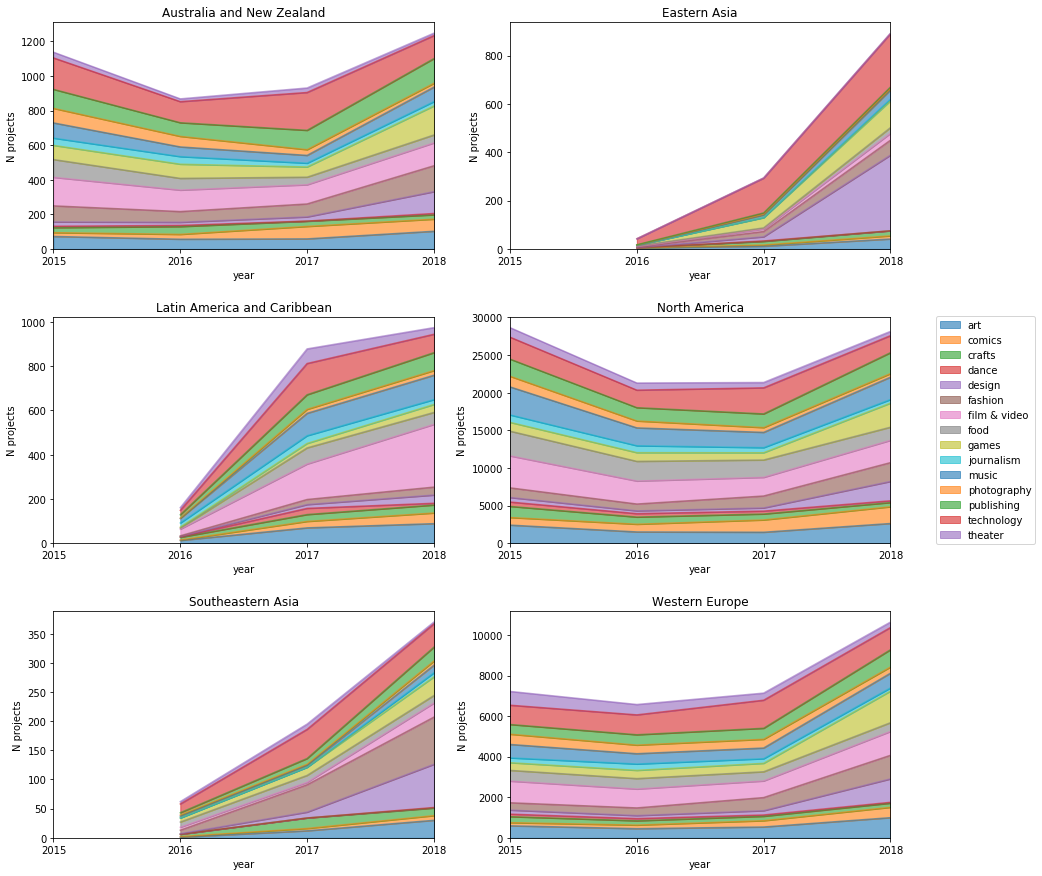
Медианная вероятность успеха кампании для каждого региона суммарно за 2015-2018г: наибольшая вероятность успеха у проектов в регионе Восточной Азии – медиана близка к 80%, наименьшая – в регионе Латинской Америки и Карибских островов – с медианой около 50%



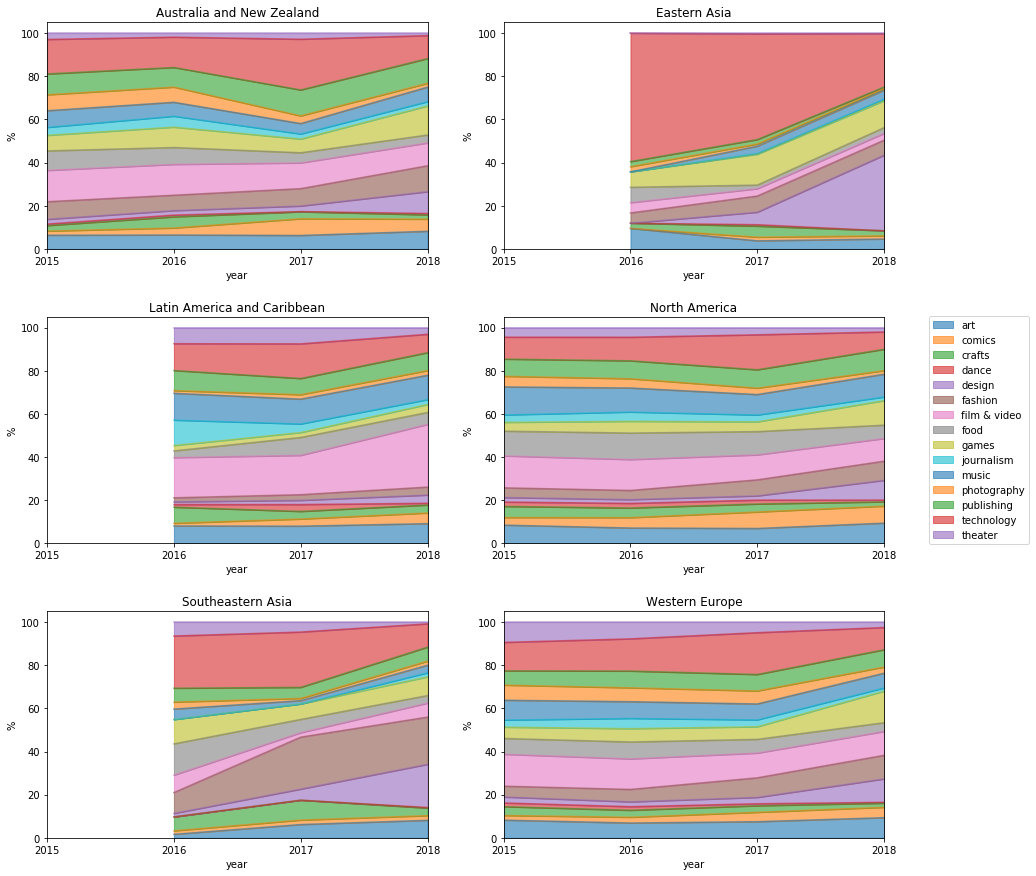
Медианная сумма вырученных средств для каждого региона суммарно за 2015-2018г: значительно больше средств собирают проекты в регионе Восточной Азии, где медиана превышает $10,000. Меньше средств собирают проекты в регионе Латинской Америки и Карибских островов – медиана около $2,000.



Проекты по категориям для каждого региона в динамике за 2015-2018: в регионах Восточной и Юго-Восточной Азии, а также Латинской Америки и Карибских островов не было проектов в кикстартере в 2015г, и за следующие года эти регионы быстро развиваются, повышая количество проектов. Среди них самый стремительный рост у Латинской Америки и Карибских островов – 1тыс проектов в 2018г. В Западной Европе наблюдается повышение общего числа проектов в 2018г почти на 3тыс проектов за год. Во всех регионах к 2018 году повышается количество дизайн проектов.



Доля проектов в каждой категории для каждого региона в динамике за 2015-2018: в Восточной Азии с 2016 по 2018 год была велика доля проектов в ИТ, в 2018 резко возросла доля дизайн проектов. В Юго-Восточной Азии схожая ситуация, однако технологичные проекты уступили свою долю индустрии моды.



Доля проектов в каждой категории для каждого региона в динамике за 2015-2018 (слева) и доля разных факторов в общем уровне счастья в среднем по каждому региону (справа). На графике не заметно общих закономерностей между факторами, формирующими уровень счастья в регионе и категориями проектов на кикстратере.

