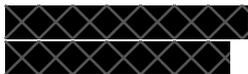


# Антон Иоков

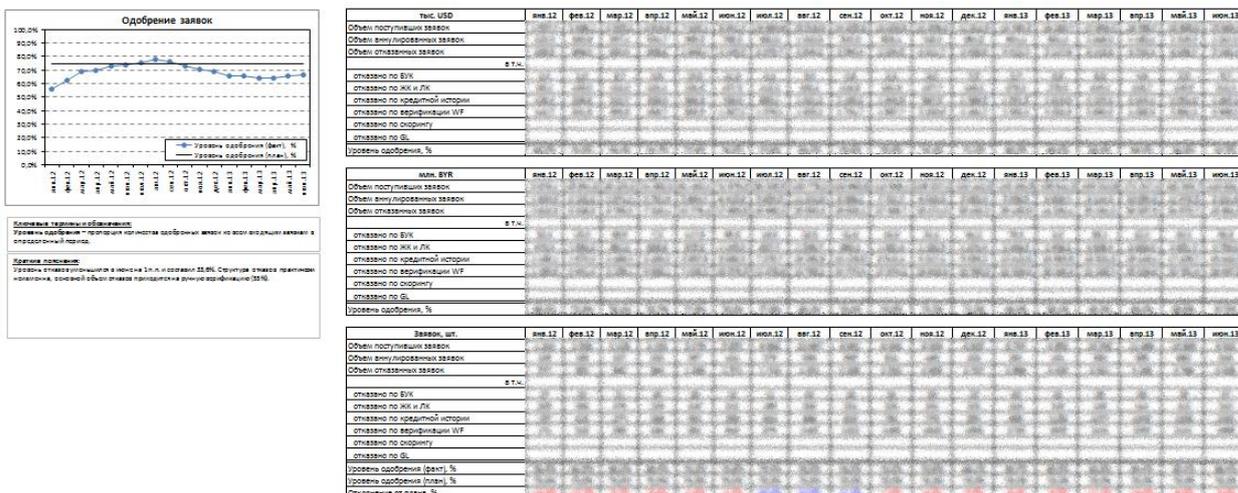


## Портфолио

В качестве портфолио я предлагаю дайджест моей работы в МТБанке. По соображениям конфиденциальности большинство данных здесь заменены или скрыты, а графики построены на специально искажённых числах.

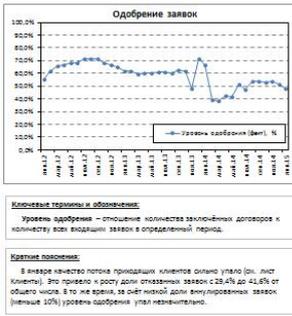
### Главный отчёт

Я пришёл в банк для того, чтобы готовить для Правления банка главный отчёт по розничным рискам, в котором полностью описывается динамика кредитного портфеля: сколько кредитов выдали, каких, кому, сколько из них вышло на просрочку и так далее. Когда мне его передавали, он выглядел вот так:



Здесь и далее — главная страница отчёта. Его полный объём — 14 таких страниц.

Из отчёта было сложно извлекать полезную информацию, и на его чтение уходило много времени. Все к этому привыкли («во всех банках так») и смирились, но меня такое положение дел не устраивало, поэтому я выяснил у членов Правления, ответы на какие вопросы они хотят увидеть в отчёте. С учётом этого я убрал некоторые разделы и логически сгруппировал по смыслу оставшуюся информацию. После чего поработал над типографикой и сеткой, избавился от лишних границ и выровнял данные:



тыс. USD	авг.13	сеп.13	окт.13	ноя.13	дек.13	январ.14	фев.14	мар.14	апр.14	май.14	июн.14	июл.14	авг.14	сеп.14	окт.14	ноя.14	дек.14	январ.15
Объём поступивших заявок	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
Объём одобренных заявок	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600
Уровень одобрения, %	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60

млн. BYR	авг.13	сеп.13	окт.13	ноя.13	дек.13	январ.14	фев.14	мар.14	апр.14	май.14	июн.14	июл.14	авг.14	сеп.14	окт.14	ноя.14	дек.14	январ.15
Объём поступивших заявок	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
Объём одобренных заявок	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600
Уровень одобрения, %	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60

Заявок, шт.	авг.13	сеп.13	окт.13	ноя.13	дек.13	январ.14	фев.14	мар.14	апр.14	май.14	июн.14	июл.14	авг.14	сеп.14	окт.14	ноя.14	дек.14	январ.15
Объём поступивших заявок	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
Объём одобренных заявок	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600
Уровень одобрения (рейт), %	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60

Несколько месяцев отчёт просуществовал в таком виде, но в какой-то момент стало понятно, что есть потребность сделать его лучше: он должен описывать не только *что* происходит, но и *почему* это происходит. К этому моменту я уже вдохновился идеями Эдварда Тафти и Дональда Нормана — и не преминул ими воспользоваться:



тыс. USD	авг.13	сеп.13	окт.13	ноя.13	дек.13	январ.14	фев.14	мар.14	апр.14	май.14	июн.14	июл.14	авг.14	сеп.14	окт.14	ноя.14	дек.14	январ.15
Объём поступивших заявок	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
Объём одобренных заявок	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600
Уровень одобрения, %	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60

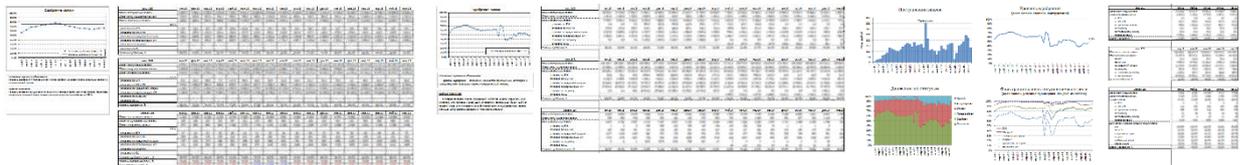
млн. BYR	авг.13	сеп.13	окт.13	ноя.13	дек.13	январ.14	фев.14	мар.14	апр.14	май.14	июн.14	июл.14	авг.14	сеп.14	окт.14	ноя.14	дек.14	январ.15
Объём поступивших заявок	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
Объём одобренных заявок	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600
Уровень одобрения, %	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60

Заявок, шт.	авг.13	сеп.13	окт.13	ноя.13	дек.13	январ.14	фев.14	мар.14	апр.14	май.14	июн.14	июл.14	авг.14	сеп.14	окт.14	ноя.14	дек.14	январ.15
Объём поступивших заявок	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000
Объём одобренных заявок	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600	600
Уровень одобрения (рейт), %	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60	60

На первый план вышли графики и диаграммы, которые помогли лучше разобраться в том, что происходит с кредитным портфелем, информация стала считываться мгновенно, а я получил положительную реакцию со стороны всех читателей отчёта.

Эволюция отчёта в small multiples:



## Скорость рассмотрения заявок

После того, как первый визуальный анализ был воспринят на ура, я получил карт-бланш на использование графики.

В день в банк приходят сотни заявок на кредит, и скоринговая система принимает решения только по части из них. Остальные отправляются на ручную проверку, и банку важно, чтобы их рассмотрели в течение некоторого времени (здесь — в течение 80 минут). Чтобы понимать, когда у нас наблюдаются проблемы со скоростью, я сделал такой хитмап:

Среднее время рассмотрения (по дням)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
	ср	чт	пт	сб	вс	пн	вт	ср	чт	пт	сб	вс	пн	вт	ср	чт	пт	сб	вс	пн	вт	ср	чт	пт	сб	вс	пн	вт	ср	чт	пт
8	2:31	0:22				1:13	0:26			1:20			0:55			4:58	3:21			0:10	0:09	1:09	0:55	1:49			1:55		2:15	1:45	1:47
9	1:16	0:44	2:10	1:56		0:21	0:27	0:25	0:28	0:51	1:48	3:45	0:38	1:09	1:56	3:36	2:22	4:51	3:31	0:41	0:29	0:34	0:47	0:55	2:42	2:33	2:29	0:57	2:08	1:32	2:21
10	1:13	0:54	0:44	2:16	0:23	0:25	0:49	0:26	1:06	1:14	1:15	3:28	0:48	1:19	2:18	3:08	2:37	4:22	2:47	1:06	0:33	0:40	0:40	0:49	1:31		2:24	0:42	1:41	1:36	2:35
11	1:30	1:04	0:40	2:14	0:21	0:57	1:13	0:56	1:24	1:05	1:09	3:10	1:51	1:16	2:11	3:16	2:09	3:23	3:53	1:04	0:59	0:38	1:04	1:08	1:30		1:59	1:14	1:55	1:32	2:11
12	1:48	1:30	1:32	1:59	0:21	1:17	1:30	1:14	1:20	1:37	1:58	3:18	2:18	2:26	3:39	3:49	2:18	3:12	2:32	1:31	1:05	1:20	1:17	1:46	2:33		1:49	2:15	2:14	2:13	2:10
13	2:07	1:46	0:53	2:06	0:27	1:27	1:40	1:07	1:46	1:42	2:45	2:50	3:02	2:28	2:47	3:04	2:09	2:28	1:13	1:36	1:17	1:25	1:50	2:06	2:18	4:39	1:35	2:06	2:17	2:38	2:22
14	1:48	1:51	0:46	1:58	0:47	1:04	1:40	1:20	1:09	1:53	2:42	2:18	2:09	3:10	2:22	1:30	1:50	0:43	1:20	1:28	1:31	0:59	1:41	2:30	1:44		1:37	2:00	2:16	2:22	2:13
15	1:43	1:49	1:04	1:13	0:33	1:11	1:38	1:02	1:31	1:41	2:12		2:58	1:41	2:38	1:19	2:12	1:24	1:08	1:08	1:28	1:09	1:06	2:16	0:37	0:48	1:14	2:08	1:48	2:13	2:06
16	1:45	1:47	0:59	0:53	0:41	1:01	1:07	0:52	1:13	1:39	1:45	0:26	1:51	1:17	2:05	1:21	2:07	1:04	0:38	1:07	1:28	1:21	1:23	1:15	0:18	1:28	1:02	1:06	1:05	1:10	1:17
17	1:28	1:07			0:31	0:48	1:09	0:50	1:18	1:30			1:03	1:11	0:46	1:36	0:45	0:26	0:24	0:48	1:36	1:15	1:03	1:06	0:30		1:14	1:05	1:42	0:59	1:35
18	1:05	0:58	0:31	0:09		0:43	0:46	0:42	1:15	1:32	1:33		0:33	1:15	0:29	1:20	0:35	0:40		0:29	0:55	0:53	0:44	0:49		1:11	0:41	1:29	1:09	0:33	
19	1:05	1:09			0:37	0:22	0:33		0:47	0:47			0:44	0:28	0:28	0:34	0:31			0:32	0:41	0:42	0:11	0:45		0:52	0:32	0:55	0:29	0:47	
20					0:31			0:43	0:21					0:21	0:11		0:22			0:22			0:22			0:25	0:10	0:21	0:09		

Вертикальная ось обозначает время суток (например, 8 часов).

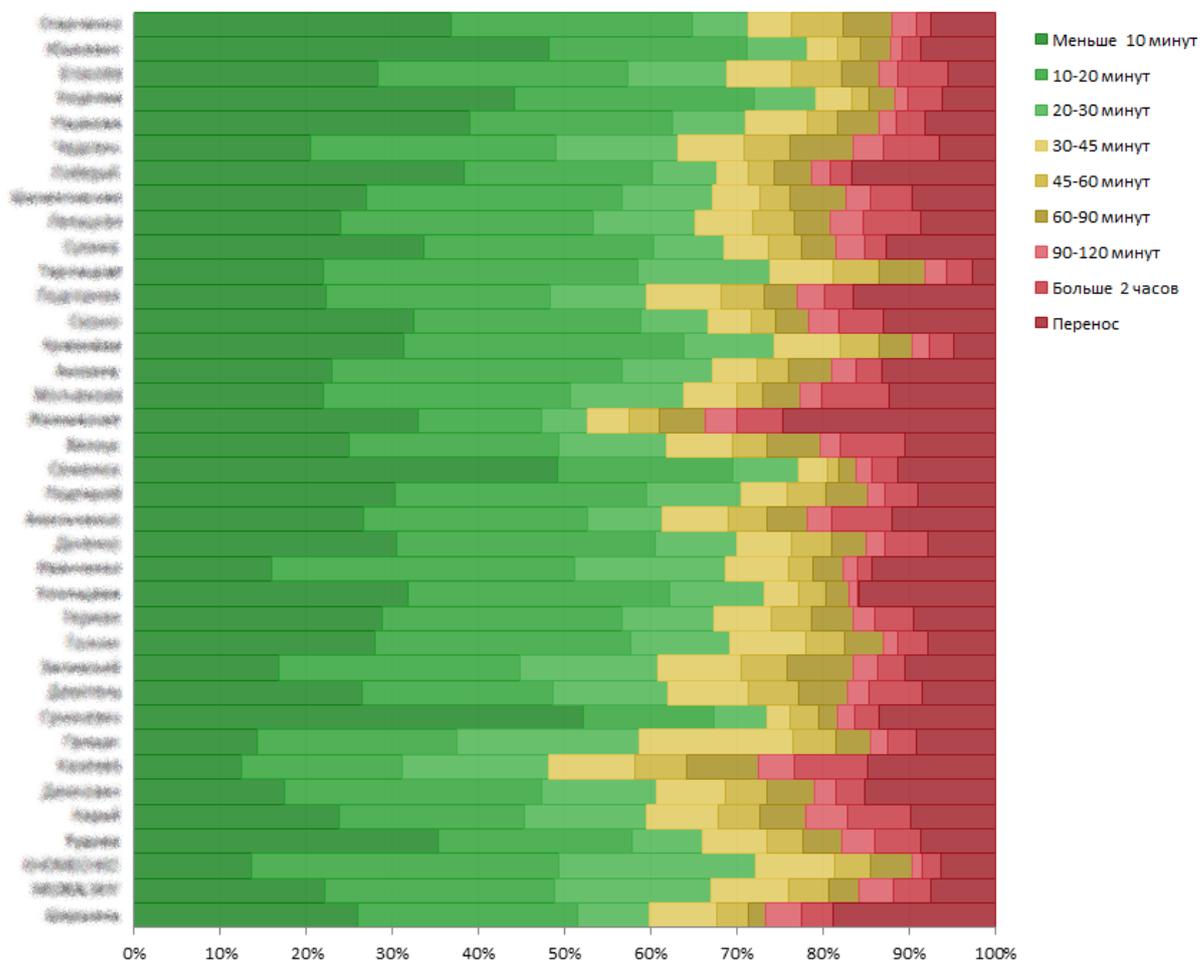
После того, как проблемные зоны стали видны невооружённым глазом, были сделаны следующие выводы:

- Нужно **перераспределить человеческий ресурс** и выводить сотрудников в другое время.
- Необходимо **запрашивать планы продаж** у дружественных подразделений, чтобы быть готовым к «валу» заявок.
- Надо **протестировать скорость выполнения запросов** к базам данных. Оказалось, что подвисшие запросы серьёзно задерживали рассмотрение заявок.

## «Быстрые» и «медленные» сотрудники

Следующей задачей стало выявление самых быстрых и самых медленных специалистов по рассмотрению заявок — верификаторов:

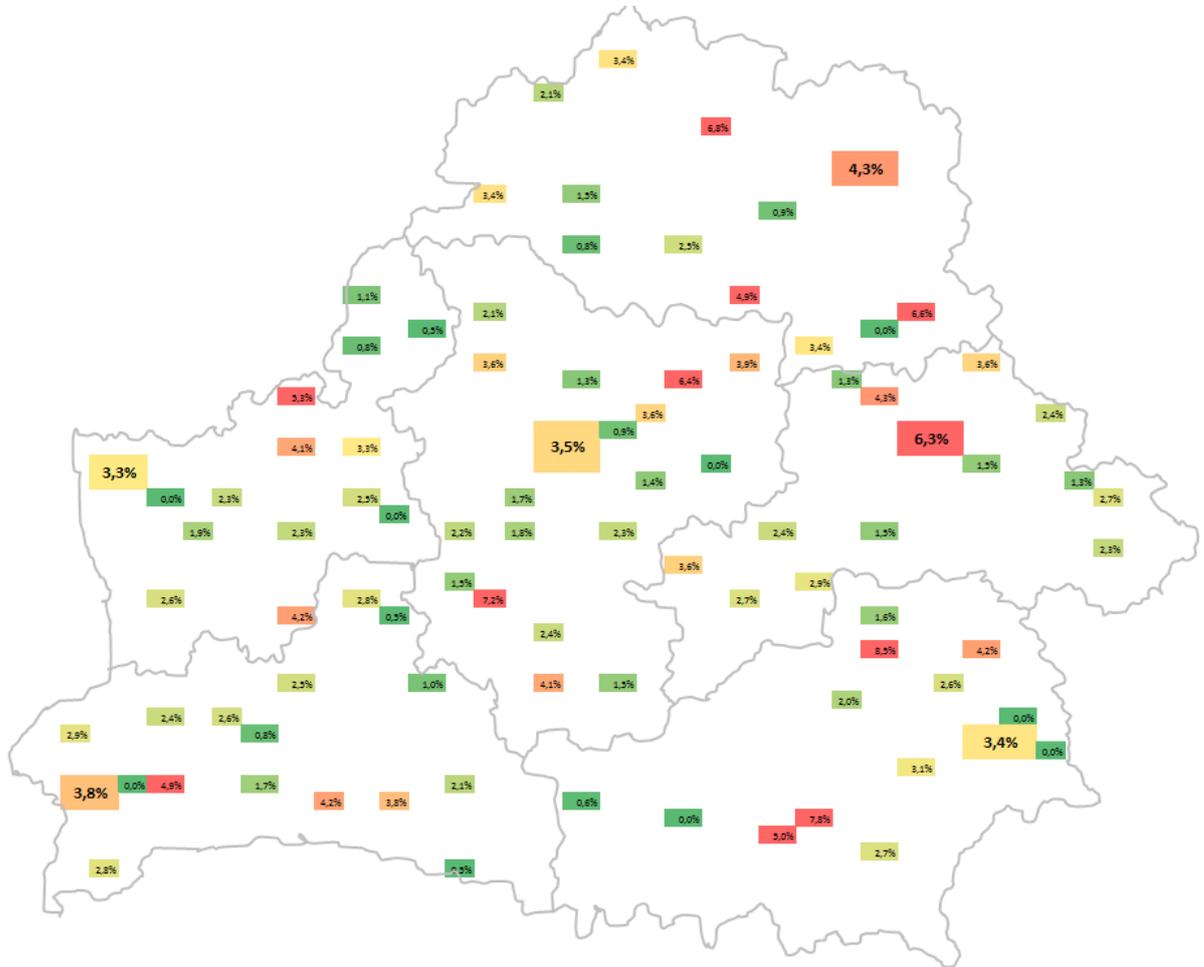
### Скорость верификации



Выводы, сделанные на основе визуализации, позволили провести точечное обучение и повысить скорость рассмотрения заявок.

## Проблемные регионы

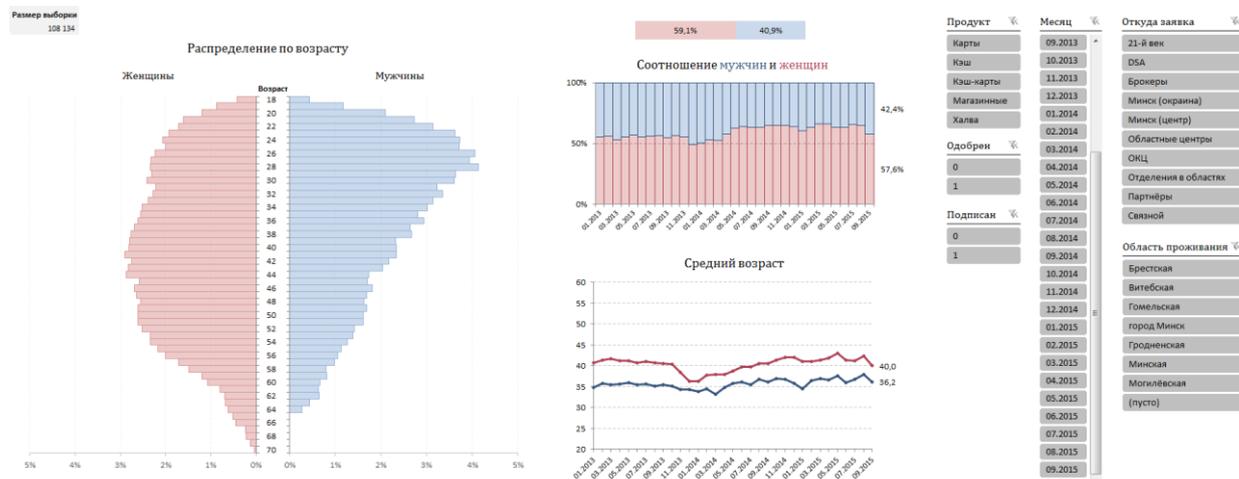
Было необходимо проверить, есть ли у банка проблемные с точки зрения выхода на просрочку регионы. Для этого нужно было нанести на карту Беларуси наши отделения и грамотно раскрасить. Но никто в банке не владел инструментами для картографической визуализации, и времени на обучение на тот момент не было, поэтому пришлось решать проблему нестандартным способом. В это время в отдел как раз ненадолго пришла практикантка, которой я поручил нарисовать в Excel контуры Беларуси и раскидать отделения по карте:



После чего условным форматированием мы добились нужного эффекта: увидели проблемные зоны и провели более глубокий анализ этих мест. После чего добавили в скоринг географию.

## Аналитика по кредитополучателям

В клиентской базе банка несколько десятков тысяч кредитополучателей, но выяснилось, что никто никогда не анализировал, что это за люди и как поток заявителей меняется со временем. Я решил исправить эту оплошность и сделал подробную визуальную аналитику:



Это вкладка «Пол и возраст», всего таких листов 9. Под каждым дашбордом есть таблицы с точными данными.

Теперь поток клиентов анализируется ежемесячно с помощью этого отчёта тремя управлениями: рисков, продаж и маркетинга.

## Обучение и консультирование

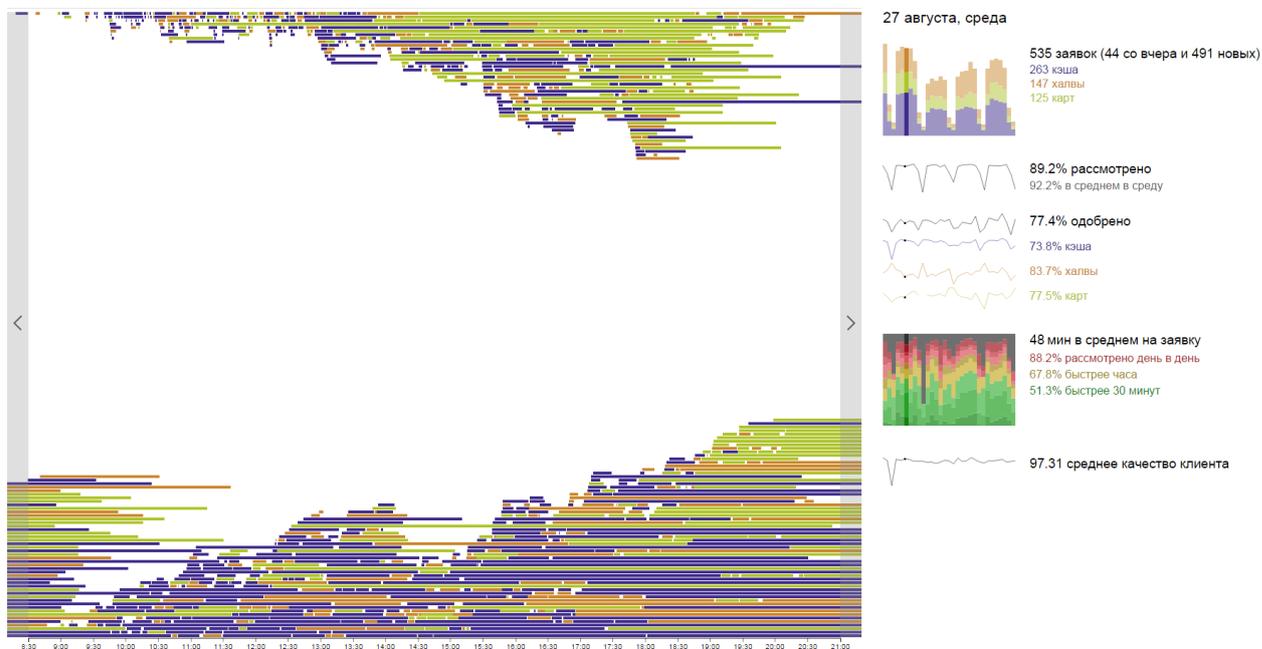
По просьбе начальника отдела я провёл обучение коллег-аналитиков базовым средствам визуализации и теперь с удовольствием наблюдаю за тем, как они используют графические средства в своей работе.

Для удобства мы выработали гайдлайны: теперь во всех отчётах управления, независимо от того, кто их готовит, используются одни и те же стандарты. «Халва» всегда цвета халвы, вспомогательные элементы не доминируют над данными, а неэффективные круговые диаграммы ушли в прошлое.

## D3.js

В какой-то момент стало понятно, что для Excel и простые графические библиотеки не подходят для решения некоторых задач, — и я стал осваивать D3.js.

Времени на спокойное освоение не было, поэтому первый же проект стал рабочим. Был необходим инструмент, позволявший осуществлять подробный мониторинг процесса верификации. Результатом стала эта визуализация:



Картинка побольше [на моём Дропбоксе](#).

Сверху на главном экране сталактитами свисают заявки, находящиеся в очереди на рассмотрение, снизу сталагмитами вырастают те, что в работе. Расположение заявок зависит от времени их поступления, длина — от времени рассмотрения, а цвет — от вида продукта.

Сразу видно, сколько и каких заявок пришло со вчерашнего дня (слева) и когда по ним было принято решение, где были основные «завалы» и с какого момента заявки стали оставаться на завтрашний день. Хорошо заметно, что во второй половине дня клиенты массово пришли за кредитными картами, и их заявки долго висели в очереди из-за низкого приоритета.

Решение разделить ожидание и рассмотрение объясняется тем, что проблемы «сталактитов» и «сталагмитов» разные по своей природе и решаются тоже совсем по-разному.

Визуализация живая: между днями можно свободно перемещаться, а заявки раскрасить не только по типу продукта, но также [по скорости их рассмотрения](#) и по конечному решению. При наведении на заявку отображается подробная информация о ней, а по клику копируется её ID.

Это первая версия проекта, и он продолжает развиваться прямо сейчас.

## Будущие проекты

В ближайшее время я выпущу новую версию последней визуализации, в которой дополню её макро- (месяц) и микро- (верификатор) уровнями.

После этого я приступлю к созданию первой картографической визуализации при помощи D3.js, чтобы проанализировать потоки клиентов.

В свободное время я делаю анализ «для души». Мне интересно, что происходит с режиссёрами после того, как они получают премию «Оскар» за лучший фильм: снимают ли они после этого более или менее качественные фильмы, растёт ли их популярность, и если да, то насколько. Раньше данные были в дефиците, а сейчас есть огромные базы Кинопоиска и IMDb, которыми я и пользуюсь для анализа.

Антон Ионов

