

# Хакатон для аналитиков от Skillfactory

## Декабрь 2022 года

### Задача №2. Анализ экспорта на основе данных ФТС

**Задача:** На основании данных ФТС РФ (база экспорта города Москвы) сформулировать гипотезы и прогнозы для изменившегося из-за санкций экспорта Москвы. В задаче 2 мы фокусируемся на глубоких гипотезах о способах решения проблем в экспорте, но в рамках решения допустимо уделять время анализу рисков и уязвимых мест.

**Входные данные:** excel таблица с данными по категориям товара, экспорту в разнообразных отраслях, список недружественных государств (вторая страница документа)

**Выходные данные:** презентация в формате:

- постановка проблемы или проблем (кратко, какая проблема в экспорте Москвы была выявлена в ходе анализа предложенных данных)
- акцент на гипотезы для закрытия проблемных мест (основываясь на анализе данных)
- какими методами анализа данных, вы пользовались
- выводы

**Критерии оценивания итогового решения (1-10 баллов):**

- корректность и точность анализа предложенных данных (обоснованность применения методов, фильтрация данных, наглядность полученных результатов)
- глубина анализа выявленных проблем(в первую очередь, в контексте создания гипотез для закрытия проблемной зоны, во вторую, в контексте анализа уязвимых мест)
- качество презентации, понятность решения
- соответствие решения запросам со стороны бизнеса

**Вариант решения задачи:**

Вы можете анализировать данные так, как считаете необходимым, применяя на практике полученные во время учебы знания. Ниже приводится поэтапный вариант решения задачи.

**На первом этапе. Подготовка и фильтрация данных.** изучить базу данных ФТС, сформировать по каждому товарному коду список ТОП-стран, импортирующих товары Москвы. Соединить базу данных со списком «Недружественных стран» и определить объем выпадающего экспорта.

**На втором этапе. Анализ данных:** определить страны, второго и третьего темпа, относящиеся к дружественным странам, и их способность выступить потребителями санкционной продукции.

**На третьем этапе. Подготовка результатов исследования, формулирование гипотез-решений.** студенты должны подготовить презентацию, показывающую какие страны, обладают наибольшим потенциалом для наращивания экспорта (перемещаются на первые места после удаления недружественных стран чаще всего), а также товарные позиции, которые нельзя заменить на основании анализа данных ФТС.

## Первые этапы работ

**Предоставленные данные:** В рамках данного хакатона МЭЦ (Московский экспортный центр) предоставил таблицу Excel, содержащей 291 947 строк. В каждой строке содержалась информация по экспорту товарных категорий (10 код ТН ВЭД) в каждую страну для каждого года с 2015 по 2021, включительно. Для нормального использования данная таблица преобразовывалась в такую, чтобы год экспорта содержался в отдельной колонке. В итоге получилась таблица с 466 976 строк - далеко не все товарные категории в отдельные страны экспортировались каждый год. Дополнительно, для удобства построения прогноза на 2022 год средствами Power BI в исходную таблицу от МЭЦ были добавлены нулевые продажи за 2022 год - это привело к увеличению количества строк основной таблицы до 759 000.

**Ожидаемая методика обработки данных:** Организаторы хакатона считали, что с такой задачей можно справиться средствами исключительно Google Table - умение работать с данными таблицами было ключевым требованием к участникам хакатона. Очевидно, что Google Table на 467 тыс. строк очень сильно тормозила, что делало работу на данном аналитическом инструменте практически не возможной. Поэтому практически все участники хакатона тут же переключились на Excel - здесь хоть и было некоторое подтормаживание, но работать было вполне возможно.

**Подход нашей команды:** Мы в своей команде решили пойти несколько дальше и использовать Power BI. Данный инструмент является идеальным средством для выбора различных информационных срезов, оперативного проведение предварительного анализа, а так же позволяет проводить экспорт выбранных срезов для обмена ими с другими членами команды. Как оказалось, ещё несколько команд так же использовали в своей работе над заданиями хакатона Power BI.

Все работы с Power BI в нашей команде проводились мной - я был единственным в нашей команде кто хотя бы более-менее был знаком с Power BI.

**Первые отчёты Power BI:** Вначале в Power BI был создан инструментарий для фильтрации и выбора любого информационного среза данных с целью экспорта данного среза для обмена с другими участниками команды для последующей обработки. См. следующую закладку...



Фильтры по кодам ТНВЭД

Товар (2 знака)

All

Товар и код ТН ВЭД (4 знака)

All

Товар и код ТН ВЭД (6 знаков)

All

Товар и код ТН ВЭД (10 знаков)

All

Фильтры по товарным категориям

Нацпроект 2021

☐ (Blank)

☐ АПК

☐ Нет

Отрасль 2021

All

Фильтры по странам

Дружественная страна

☐ Enemy

☐ Friend

Регион РЭЦ

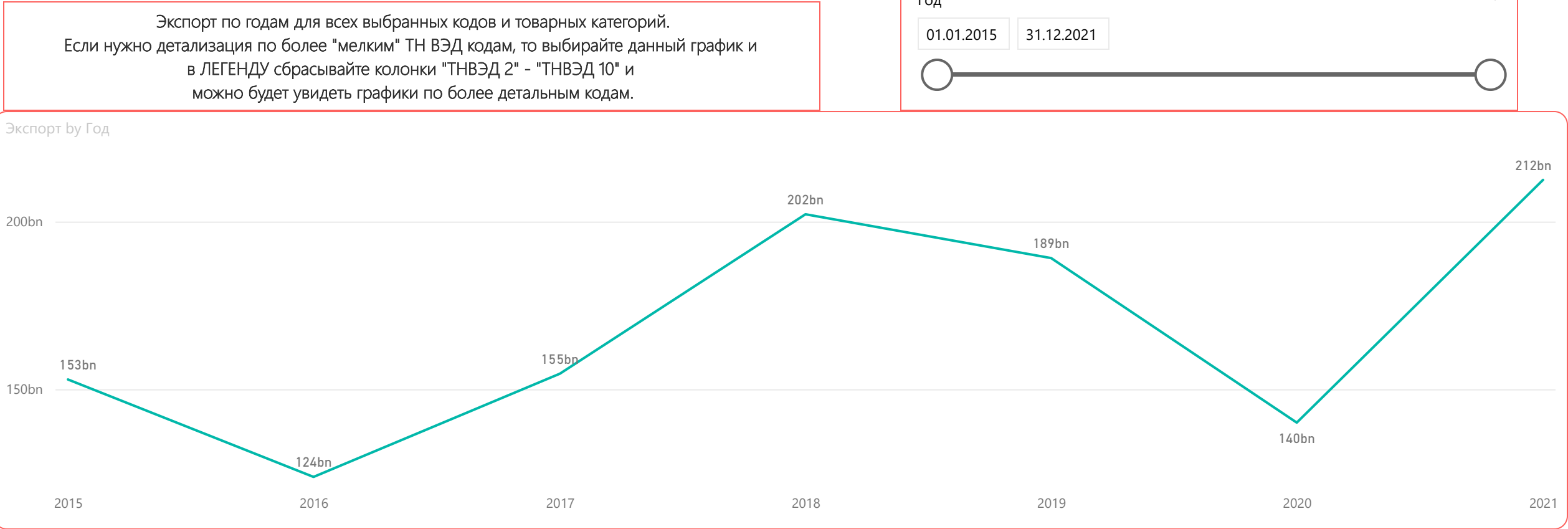
All

Регионы (РЭЦ) и Союзы

All

Страна

All



## Последующие этапы работ

После разработки с помощью Power BI механизма выгрузки любых информационных срезов для всех членов команды, было принято решение:

"А почему бы не воспользоваться возможностями Power BI для проведения анализа исторических данных за период 2015-2021 года, а так же, расчёта предполагаемого экспорта на 2022 год?"

Сказано - сделано.

Результат - на следующих 2 закладках.

Пара слов о модели расчёта экспорта на 2022 год.

Был использован следующий алгоритм:

1. Экспорт в НЕ дружественные страны (Enemy) приравнивался к нулю - последствия санкций.
2. Экспорт в дружественные страны (Friend), для компенсации выпадающего экспорта, может быть увеличен на определённое количество процентов, задаваемое в соответствующем поле. Данный процент можно задать от 0% до 100% - другими словами, увеличение экспорта в дружественные страны возможно до 2 раз.

Следует отметить, что технология "dashboard" предлагает прекрасные возможности для получения ответов на все вопросы в рамках задачи №2 данного хакатона.

См. следующие 2 закладки...



Фильтры по кодам ТНВЭД

Товар (2 знака)

All

Товар и код ТН ВЭД (4 знака)

All

Товар и код ТН ВЭД (6 знаков)

All

Товар и код ТН ВЭД (10 знаков)

All

Фильтры по товарным категориям

Нацпроект 2021

☐ (Blank)

☐ АПК

☐ Нет

Отрасль 2021

All

Фильтры по странам

Регион РЭЦ

All

Регионы (РЭЦ) и Союзы

All

Страна

All

Дружественная страна

☐ Enemy

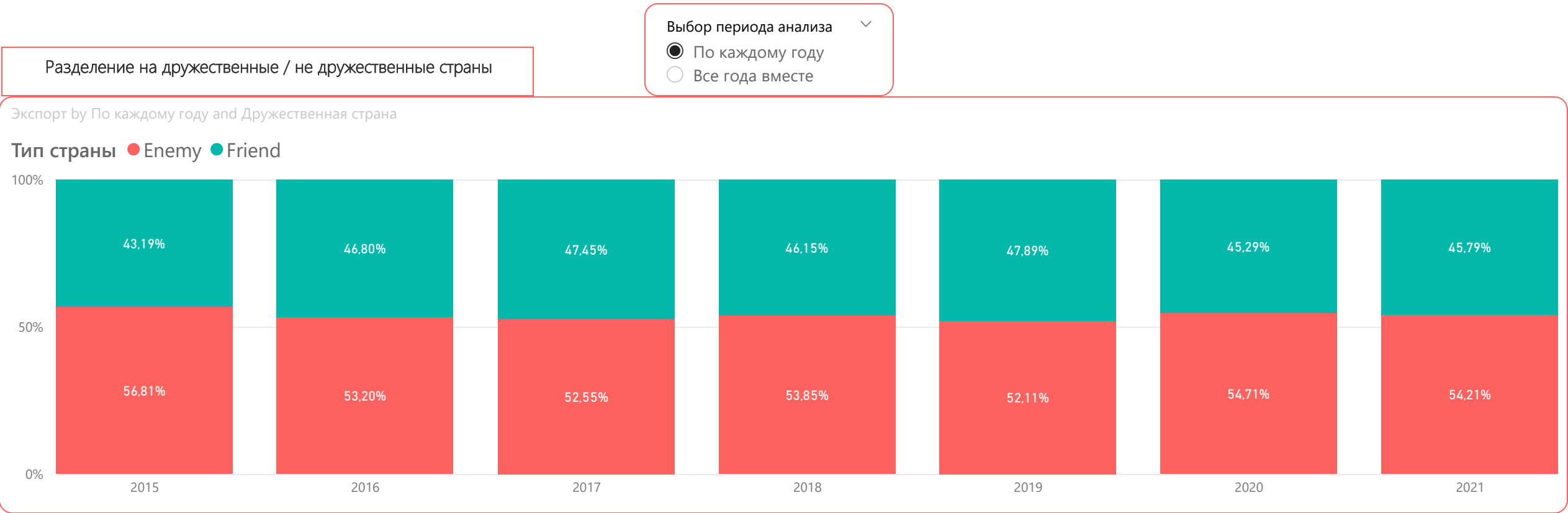
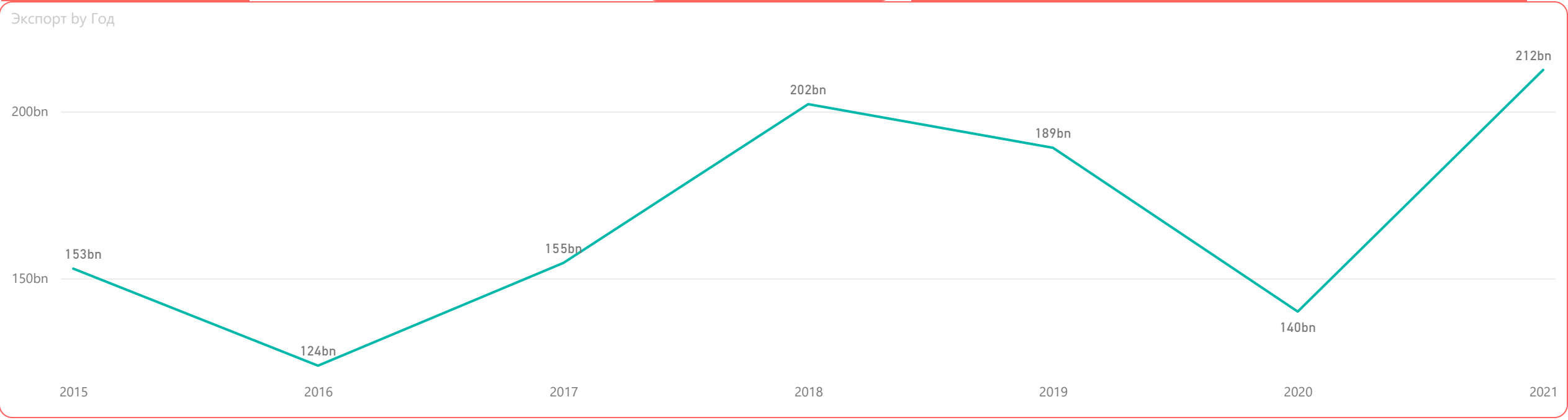
☐ Friend

Год

01.01.2015

31.12.2021

Суммарный экспорт



Детализации по ТН ВЭД коду

Кол-во цифр в коде ТН ВЭД

Выводимый параметр

ТНВЭД 4	2015	2016	2017	2018	2019	2020	2021	Total
0101	\$449 080	\$447 850	\$145 875	\$58 916	\$123 005	\$37 332	\$276 339	\$1 497 007
0102		\$490			\$32 737	\$166	\$220 200	\$220 893
0104				\$42 955			\$587 185	\$629 140
0105	\$121 726							\$121 726
0106	\$1 091 517	\$1 418 200	\$1 511 858	\$433 112	\$640 726	\$87 909	\$181 998	\$5 085 210
02					\$8 028 734	\$11 318 456	\$73 768	\$19 355 948
0201	\$325	\$90 195	\$35 577	\$67 289	\$184 031	\$171 322	\$835 656	\$1 484 295
0202	\$105 741	\$237 849	\$336 498	\$1 346 334	\$7 787 659	\$12 810 514	\$18 802 169	\$24 416 614
0203	\$84 038	\$519 083	\$185 229	\$99 806	\$568 015	\$7 265 305	\$2 455 225	\$9 519 691
0204	\$229	\$118 866	\$812 033	\$8 502 619	\$16 553 474	\$99 353	\$77 328	\$23 955 002
0205					\$14 838	\$3 881	\$7 046	\$25 765
0206	\$4 956 088	\$12 672 572	\$19 051 302	\$16 773 725	\$8 665 302	\$8 809 518	\$5 060 680	\$69 697 187
0207	\$2 385 039	\$8 907 255	\$19 329 712	\$28 145 499	\$29 419 253	\$40 751 713	\$42 251 940	\$191 641 411
0208	\$6 553	\$63 769	\$105 489	\$33 015	\$46 498	\$26 991	\$29 788	\$199 103
0209	\$6	\$13 048	\$292 582	\$300 617	\$378 702	\$369 432	\$548 051	\$1 562 438
0210	\$49 722	\$81 625	\$22 895	\$19 351	\$42 207	\$74 700	\$58 015	\$287 415
Total	\$66 046 966 708	\$57 952 944 319	\$73 369 838 996	\$93 300 293 940	\$90 565 128 087	\$63 423 332 506	\$97 290 827 294	\$541 940 299

Если среди 4, 6 и 10 значных кодов стоит 2 значный ТН ВЭД код, то страной получателем данного экспорта является страна "NNN-НЕИЗВЕСТНАЯ СТРАНА". В рамках данного анализа НЕТ основания удалять данные записи по экспорту из проводимого анализа. Если Вы хотите удалить данную страну из проводимого анализа, то я выше в слайсере по странам отмените галочку рядом со страной "NNN-НЕИЗВЕСТНАЯ СТРАНА".

Дополнительные параметры по кодам

Выбор периода анализа для таблицы

Выводимые параметры

ТНВЭД 4	2015	2016	2017	2018	2019	2020	2021	Total
0101	91%	79%	85%	60%	64%	15%	77%	72%
0102		62%			100%	100%	100%	100%
0104				100%			100%	100%
0105	100%							100%
0106	98%	93%	100%	99%	90%	32%	26%	86%
02					100%	100%	100%	100%
0201	100%	100%	100%	100%	100%	100%	100%	100%
0202	100%	100%	100%	100%	100%	100%	100%	100%
0203	30%	89%	4%	5%	100%	100%	100%	62%
0204	100%	100%	100%	100%	100%	100%	100%	100%
0205					100%	100%	100%	39%
0206	100%	100%	100%	100%	100%	100%	100%	100%
0207	96%	92%	100%	100%	100%	100%	100%	99%
0208	100%	100%	100%	100%	100%	100%	100%	100%
0209	100%	96%	100%	79%	100%	100%	100%	96%
0210	97%	91%	100%	100%	100%	100%	100%	97%
03							100%	100%
0301		100%	100%		99%	100%	100%	100%
0302	100%	99%	100%	100%	97%	100%	100%	100%
0303	82%	71%	58%	73%	66%	75%	77%	71%
0304	47%	83%	89%	89%	81%	80%	87%	84%
Total	43%	47%	47%	46%	48%	45%	46%	46%

Используются следующие расчётные параметры:

CR3 - Индекс концентрации рынка, доля 3 стран с наибольшей долей экспорта. 0-45% - низкая оценка (зелёный цвет), 45-70% - умеренная оценка (жёлтый цвет), 70-100% - высокая оценка (красный цвет).

НН1 - Индекс монополизации Херфиндаля-Хиршмана, сумма квадратов процентных долей рынка ВСЕХ стран-импортёров. Максимальное значение - 10 000, когда одна страна участница с долей 100%. 0-1 000 - низкая оценка (зелёный цвет), 1 000-2 000 - умеренная оценка (жёлтый цвет), 2 000-10 000 - высокая оценка (красный цвет).

Доля Friends - доля экспорта, которая приходится на дружественные страны. Оценка возможности замены экспорта в НЕ дружественные страны увеличением экспорта в дружественные. 80-100% - низкая угроза (зелёный цвет), 50-80% - умеренная угроза (жёлтый цвет), 0-50% - высокая угроза (красный цвет).

Фильтры по кодам ТНВЭД

Товар (2 знака)

All

Товар и код ТН ВЭД (4 знака)

All

Товар и код ТН ВЭД (6 знаков)

All

Товар и код ТН ВЭД (10 знаков)

All

Фильтры по товарным категориям

Нацпроект 2021

☐ (Blank)

☐ АПК

☐ Нет

Отрасль 2021

All

Фильтры по странам

Регион РЭЦ

All

Регионы (РЭЦ) и Союзы

All

Страна

All

% увеличения экспорта в Friend страны

30

Дружественная страна

☐ Enemy

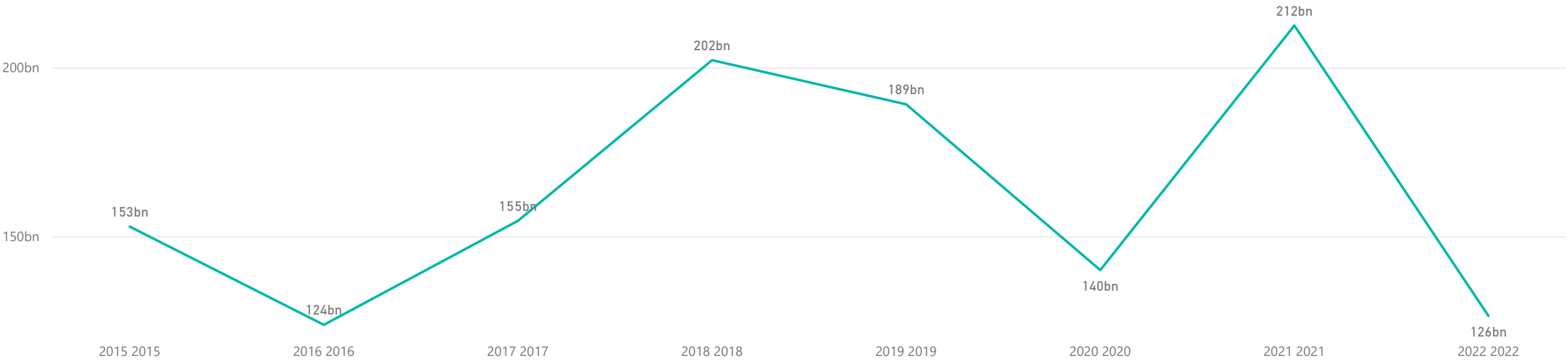
☐ Friend

Год

01.01.201501.01.2022

Суммарный экспорт

2015-2021, Плюс 2022 by DateYear and Год



Если среди 4, 6 и 10 значных кодов стоит 2 значный ТН ВЭД код, то страной получателем данного экспорта является страна "NNN-НЕИЗВЕСТНАЯ СТРАНА".

В рамках данного анализа НЕТ основания удалять данные записи по экспорту из проводимого анализа.

Если Вы хотите удалить данную страну из проводимого анализа, то я выше в слайсере по странам отмените галочку рядом со страной "NNN-НЕИЗВЕСТНАЯ СТРАНА".

Детализаций по ТН ВЭД коду

Кол-во цифр в коде ТН ВЭД

ТНВЭД 2

ТНВЭД 4

ТНВЭД 6

ТН ВЭД 10

ТНВЭД 2	2015	2016	2017	2018	2019	2020	2021	2022
01	1 729 925	2 094 775	1 690 663	577 805	937 467	527 101	1 857 815	1 645 439
02	7 884 592	23 525 259	44 926 294	57 325 066	71 708 460	81 757 419	70 226 532	91 259 565
03	19 241 158	45 166 339	47 612 169	53 887 236	55 484 562	54 085 808	61 320 226	63 857 679
04	18 847 726	19 547 063	27 420 079	24 876 270	21 801 817	28 399 782	42 642 828	51 799 922
05	13 979 473	13 146 143	17 989 229	16 050 968	11 409 533	8 969 726	7 768 511	3 448 090
06	225 027	203 023	185 422	174 315	466 857	559 858	1 161 508	1 404 013
07	46 362 456	45 160 415	53 362 189	33 576 044	37 765 076	51 800 735	115 316 551	133 396 562
08	6 802 877	11 440 249	21 797 797	12 784 471	16 043 150	8 822 521	19 344 588	9 443 127
09	15 694 245	23 258 785	28 660 400	33 445 674	29 790 033	32 876 263	41 434 041	48 812 056
10	1 755 264 105	1 437 438 980	1 973 793 849	2 265 454 482	1 593 248 348	2 334 192 752	2 041 882 812	2 467 703 188
11	24 206 197	20 196 048	28 426 709	25 418 152	41 813 963	37 717 171	47 437 736	45 264 007
12	7 781 126	11 595 406	21 721 516	18 534 411	59 616 827	134 683 736	92 761 857	113 413 052
13	390 616	396 664	913 720	1 132 729	1 450 894	1 201 786	938 663	1 068 374
14	1 428 268	580 763	1 136 859	1 109 011	1 719 281	910 620	498 211	292 859
15	56 675 551	56 722 293	98 052 166	120 430 889	151 276 988	135 305 698	146 543 791	166 263 140
16	31 097 583	33 520 914	49 896 992	56 250 093	65 977 842	68 398 904	87 601 443	104 656 961
17	27 987 086	46 757 781	103 521 070	129 560 492	128 769 595	146 962 799	176 258 224	200 512 986
Total	152 906 850 946	123 832 504 016	154 622 865 788	202 183 392 022	189 111 221 843	140 033 375 162	212 449 730 097	126 478 075 482

Страна	2015	2016	2017	2018	2019	2020	2021	2022
CN-КИТАЙ	14 260 802 912	13 526 754 597	17 491 685 208	29 609 920 397	29 660 377 564	21 746 216 249	29 887 312 270	38 853 505 951
NL-НИДЕРЛАНДЫ	16 443 164 675	11 536 355 395	14 189 104 857	17 302 164 346	16 883 332 350	8 702 525 278	15 948 866 221	20 733 526 088
TR-ТУРЦИЯ	10 445 266 313	6 430 449 197	9 171 562 965	11 487 571 990	12 666 057 775	6 804 567 637	12 364 355 125	16 073 661 663
BY-БЕЛАРУСЬ	7 954 134 642	7 255 556 585	9 004 235 300	10 078 063 410	9 801 827 413	7 133 865 588	11 190 864 505	14 548 123 856
KZ-КАЗАХСТАН	1 876 455 671	2 172 041 613	3 194 658 663	3 242 605 407	3 592 862 487	3 225 671 253	3 898 277 975	5 067 761 367
IN-ИНДИЯ	1 623 240 673	1 948 691 849	2 528 563 652	2 601 766 911	2 548 652 973	1 304 258 724	2 697 041 752	3 506 154 277
DZ-АЛЖИР	1 430 393 203	3 119 681 389	4 468 245 241	2 869 267 326	1 317 024 382	1 888 309 155	2 369 174 908	3 079 927 381
BE-БЕЛЬГИЯ	450 861 218	488 828 319	725 028 751	1 580 549 497	1 025 307 759	762 018 343	1 399 364 078	1 819 173 302
NNN-НЕИЗВЕСТНАЯ СТРАНА	2 905 823	2 271 623	3 878 906	4 132 518	157 136 100	627 533 248	1 263 300 244	1 642 290 317
AE-ОБЪЕДИНЕННЫЕ АРАБСКИЕ ЭМИРАТЫ	394 681 235	172 593 370	361 906 580	249 763 124	210 470 703	425 839 473	1 182 136 855	1 536 777 912
HK-ГОНКОНГ	447 300 670	297 870 194	208 680 888	302 009 746	407 944 383	1 716 373 469	1 148 980 408	1 493 674 531
MD-МОЛДОВА, РЕСПУБЛИКА	770 093 510	623 008 041	530 713 268	740 250 438	774 883 279	533 509 605	1 145 361 407	1 488 969 830
AM-АРМЕНИЯ	675 514 749	580 626 047	750 789 444	768 918 945	1 210 013 517	828 888 023	989 501 486	1 286 351 932
SA-САУДОВСКАЯ АРАВИЯ	216 150 890	29 260 342	340 795 733	239 626 258	392 711 660	534 087 148	930 139 329	1 209 181 127
VN-ВЬЕТНАМ	775 088 902	418 704 499	1 322 456 787	497 741 881	73 965 285	105 555 641	850 207 983	1 105 270 378
RS-СЕРБИЯ	469 077 299	359 572 390	514 568 365	572 342 474	904 818 648	316 319 018	824 521 415	1 071 877 840
MY-МАЛАЙЗИЯ	77 367 055	292 472 744	60 943 551	398 723 568	605 019 707	608 697 024	795 198 287	1 033 757 773
Total	152 906 850 946	123 832 504 016	154 622 865 788	202 183 392 022	189 111 221 843	140 033 375 162	212 449 730 097	126 478 075 482