

Проект «Реформа управління на сході України II»,
що виконується за дорученням Уряду Німеччини



Виконавець:



Рекомендації по роботі з просторовими даними (для неспеціалістів)

Київ, 2019

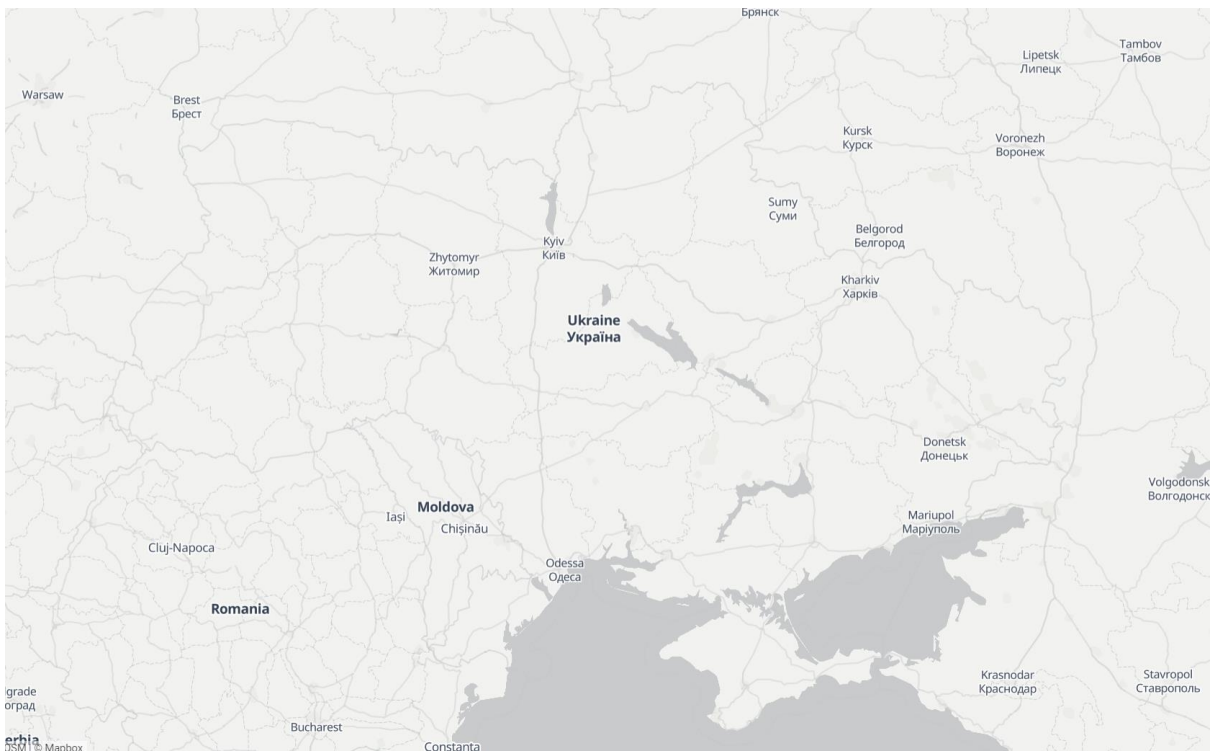
Аналіз просторових даних

Органи місцевого самоврядування поступово приходять до розуміння необхідності переходу від етапу збору та накопичення різних просторових даних до їх вивчення для отримання корисної інформації. Навіть якщо до реального аналізу зараз дійшла невелика частка організацій (більшою мірою в бізнесі), то розуміння його необхідності вже поглинула всі сфери. Однією з головних проблем початку роботи з програмним забезпеченням для дослідження просторових даних є неповне розуміння як із звичних інструментів збору даних перейти до даних специфічного формату. У цьому документі ми коротко опишемо основні підходи до організації просторових даних, необхідних для роботи зі спеціалізованим програмним забезпеченням.

Що таке просторові дані? Серед фахівців часто поширений підхід до розуміння даних оснований на геометрії:

1. Точкові дані: ЦНАПи, відділення банкоматів, місця фіксації злочинів, місця здійснення дзвінків чи відправлення повідомлень, фотографування, розміщення МАФів з кавою.
2. Лінійні дані: шляхи доріг, невеликі річки, рух користувачів, кораблів.
3. Полігональні дані: межі громад, міст та країн, озера та великі річки, зони обслуговування.

80% просторових даних мають точкове вираження, тому зосередимося на ньому. Зазвичай програма очікує дані у вигляді просторових координат, з яких і почнемо роботу.



Просторові дані. Як їх правильно вказати?

Просторові координати зазвичай передаються у вигляді широти та довготи.

Підхід 1. Збереження даних у вигляді градусів, хвилин, секунд. Цю інформацію можна проглянути клікнувши на будь яке місце на мапі Google та проглянувши бокову панель: [50°27'31.9"N 30°31'05.2"E](#). Часто такий підхід використовується в офіційних документах, наприклад [щодо меж тимчасово окупованих територій у Донецькій та Луганській областях](#). Використання спеціальних символів ° ' " є недоліком підходу, а також додаткова складність полягає у необхідності розуміння східної-західної та північної-південної півкуль. Хоч це і шкільна програма, та часто це може бути зайвим клопотом.

Програмне забезпечення часто «не зрозуміє», що це геометрія саме через використання спеціальних символів. За наявності даних у такому форматі рекомендується їх конвертувати у наступний описаний формат.

Підхід 2. Збереження даних у вигляді десяткових дрібних значень градусів. Все ті ж градуси, але без спеціальних символів та необхідності розуміння півкуль. Таку інформацію також можна знайти навіть у пошуку Google, для чого треба ввести наступний запит: [50.458861, 30.518111](#).

Будь яке сучасне програмне забезпечення «вміє» стабільно працювати з такими даними. Тому саме такий вигляд рекомендується для збереження та подальшого використання географічних координат.

Робота з таблицями

Розуміючи як правильно вказувати координати, можна переходити до підготовки та збору даних, які нас цікавлять. Інструментом для виконання цієї задачі можуть бути звичайні електронні таблиці.

Розглянемо основні правила оформлення таблиць у Excel, Open/Libre Office та Google таблицях:

- ніколи не об'єднувати осередки таблиці – всі колонки мають залишатися самостійними, навіть якщо дані можна логічно згрупувати;
- не пропускати стрічки в таблиці для виділення якихось категорій об'єктів;
- не пропускати рядки зверху документа;
- не вказувати назви таблиць зверху самої таблиці, для цього можна просто змінити назву таблиці у відповідній вкладці;
- використовувати стислі закодовані назви для колонок – їх пояснення потребують значно більше місця, ніж варто вписувати в назву. Для цього можна створити окремий документ або вкладку;

- не використовувати символів коми, крапки з комою, табуляції, двокрапки у назвах колонок – програма має чітко зрозуміти де назви, а де символи всередині назв, коли вона використовує ці символи;
- в кожну колонку вносити лише один тип даних – не варто вказувати тип валюти в колонку з сумою;
- використовувати формат даних boolean в разі виникнення задачі вказати на наявність або відсутність чогось. Не рекомендується використовувати значення типу «+» і «-», «1» і «04», «Так» і «Ні», «так» і «ні», «наявне» і «немає». Рекомендовано «true» та «false»;
- номери телефонів записувати в єдиному форматі. Однаковий формат введення всіх однотипних даних важливіший за конкретний обраний стандарт;
- не скорочувати кількість знаків після коми у даних координат – це нормально, що там їх може бути багато.

_id	№ п/п	Назва	Адреса	П.І.Б. ке...	Телефо...	E mail
23	№п/п	Назва	Адреса	П.І.Б. ке...	Телефо...	E mail
24	№п/п	Назва	Адреса	П.І.Б. ке...	Телефо...	E mail
25	1.	Покровський район Криворізька ...	50029, вул. Ракітіна (Ногіна), 22	Ткаченк...	95-92-34	kss71@u...
26						
27		ЗАКЛАДИ ДОШКІЛЬНОЇ ОСВІТИ				
28						
29	№ п/п	Назва	Адреса	П.І.Б. ке...	телефон	E mail
30	42.	Покровський район Комунальний ...	50015, вул. Кропивницького, 11	Басацьк...	95-41-48	kdnz25@...
31	43.	Комунальний заклад «Дошкільний...»	50014, вул. Тихвінська, 1Е	Паученк...	50-44-81	kdnzk26...
32	44.	Комунальний заклад «Дошкільний...»	50012, вул. Сеченова, 47	Федоро...	66-01-20	fedorova...
66	№ п/п	Назва	Адреса	П.І.Б. ке...	Телефо...	E mail
67		Покровський район				
33	45.	Комунальний заклад «Дошкільний...»	50042, вул. Курчатова, 6	Міщук Л...	096-012...	kkdnz31...
34	46.	Комунальний заклад «Дошкільний...»	50106, мкрн.5-й Зарічний, 10-А	Мовчан ...	95-32-97	kzdnz33...
35	47.	Комунальний заклад «Дошкільний...»	50014, вул. Руденка, 3	Кузьма ...	440-33-69	kkdnz60...
36	48.	Комунальний заклад «Дошкільний...»	50081, мкрн.5-й Зарічний	Сівоха ...	95-63-87	dnzk62...
37	49.	Комунальний заклад «Дошкільний...»	50081, мкрн.5-й Зарічний 46А	Єсіна Ір...	95-76-51	kkdnz 65...
38	50.	Комунальний заклад «Дошкільний...»	50078, мкрн. Індустріальний, 69-а	Руда На...	442-63-01	kkdnz71...
39	51.	Комунальний заклад «Дошкільний...»	50106, мкрн.4-й Зарічний	в. о. Стр...	95-30-43	kdnz76@...
40	52.	Комунальний заклад «Дошкільний...»	50053, провулок. Буглерова, 12	Безпала...	51-02-92	kdnz104...
41	53.	Комунальний заклад «Дошкільний...»	50063, вул. Едуарда Фукса (Тухачевсько...	Хайнак ...	66-23-29	dnz106d...
42	54.	Комунальний заклад «Дошкільний...»	50061, вул. Ракітіна (Ногіна), 18А	Сахальо...	95-86-09	kdnzvet...
43	55	Комунальний заклад «Дошкільний...»	50081, вул. Бажитова, 10-22	Ушачко...	95-36-81	dnz118...

Приклад невдалого формування таблиці з даними: з'єднано кілька таблиць. Назви районів та тип закладу, можна було вказати як додатковий атрибут, зберігши цілісність даних для машинної обробки. В деяких осередках вказано декілька номерів телефону, що також ускладнює подальше опрацювання. Назви колонок вказано українською з використанням спеціальних символів. [Джерело](#)

Маючи таблицю з правильно внесеними даними, можна їх аналізувати. Зазвичай при використанні таких даних в тому чи іншому програмному забезпеченні, система автоматично зрозуміє, що за тип даних вказаний у колонках, однак наші координати фактично мають тип даних число з комою. Як саме дати системі зрозуміти, що ми маємо географічні координати?

Існує кілька неформальних «правил» щодо того, як називати колонки, щоб система зрозуміла, що ми хочемо створити карту. Для назви колонок широти та довготи рекомендується використовувати наступні значення:

- «longitude» і «latitude» – довгота та широта англійською;
- «lon» та «lat» – теж саме скорочено;
- «lg» та «lt» – теж саме скорочено;

- «х» та «у» – теж саме, але більш математично представлено.

Правильно названі колонки зазвичай дозволяють без перешкод почати роботу з картографічними сервісами.

_id	identifier	address...	address...	address...	streetA...	locationD...	latitude	longitude	itemSpe...	itemQua...	itemQua...	datePla...	planted...	f
1	2018-001	Україна	Кіровог...	Кропивн...	бульвар...	вздовж в...	48.5048...	32.2225...	Горобина	2	2	2018-04...	ГРОМА...	1
2	2018-002	Україна	Кіровог...	Кропивн...	бульвар...	вздовж в...	48.5048...	32.222593	Клен го...	4	4	2018-04...	ГРОМА...	1
3	2018-003	Україна	Кіровог...	Кропивн...	бульвар...	вздовж в...	48.5048...	32.2225...	Береза	4	4	2018-04...	ГРОМА...	1
11	2018-011	Україна	Кіровог...	Кропивн...	бульвар...	біля буд.к...	48.5159...	32.275961	Туя	10	10	2018-04...	КІРОВО...	2
8	2018-008	Україна	Кіровог...	Кропивн...	вул.Вок...	вздовж в...	48.5117...	32.2474...	Липа се...	11	11	2018-04...	КОЛЕКТ...	2
5	2018-005	Україна	Кіровог...	Кропивн...	вул.Мал...	вздовж в...	48.5193...	32.2803...	Липа се...	37	37	2018-04...	КОЛЕКТ...	2
14	2018-014	Україна	Кіровог...	Кропивн...	вул.Оле...	біля буд.к...	48.5192...	32.2583...	Шовков...	1	1	2018-04...	КРОПИ...	7
13	2018-013	Україна	Кіровог...	Кропивн...	вул.Оле...	біля буд.к...	48.5192...	32.258386	Верба бі...	2	2	2018-04...	КРОПИ...	7
4	2018-004	Україна	Кіровог...	Кропивн...	вул.Пол...	вздовж в...	48.521781	32.2907...	Липа се...	70	70	2018-04...	КОЛЕКТ...	2
6	2018-006	Україна	Кіровог...	Кропивн...	вул.Соб...	вздовж в...	48.5013...	32.251125	Липа се...	27	27	2018-04...	КОЛЕКТ...	2
10	2018-010	Україна	Кіровог...	Кропивн...	вул.Шмі...	біля прив...	48.528445	32.2866...	Липа се...	7	7	2018-04...	КОЛЕКТ...	2
9	2018-009	Україна	Кіровог...	Кропивн...	набере...	null	47.8600...	32.269251	Туя	32	32	2018-04...	КОЛЕКТ...	2
7	2018-007	Україна	Кіровог...	Кропивн...	набере...	null	47.8600...	32.2692...	Липа се...	30	30	2018-04...	КОЛЕКТ...	2
12	2018-012	Україна	Кіровог...	Кропивн...	площа Д...	біля буд.к...	48.5201...	32.2835	Береза	10	10	2018-04...	ДЕРЖА...	1

Приклад більш вдалого збереження даних. Адреси збережено окремими колонками, широта і довгота збережені гарно. Дата та час збережено згідно стандарту ISO 8601. [Джерело](#)

Геокодування

Адреси використовуються набагато частіше ніж координати. Зазвичай локація об'єктів вказується у вигляді вказівки на країну, місто, вулицю та номер будинку. Є два підходи до збереження такої інформації:

Підхід 1. Єдина колонка. Вся адреса пишеться в одній колонці: "Київ, Андріївський узвіз, 23, 02000". Має свої переваги та недоліки, але в цілому приймається більшістю сервісів. Порада – називати цю колонку «address».

Підхід 2. Окремі колонки. Адреса розбивається на окремі колонки: країна, місто, район, вулиця і т.д. Залежно від країни це може відрізнятись. Зазвичай програмам легше сприймати інформацію подану в такому вигляді, і точність виявляється вищою. Обираючи цей підхід, рекомендується використовувати наступні назви колонок:

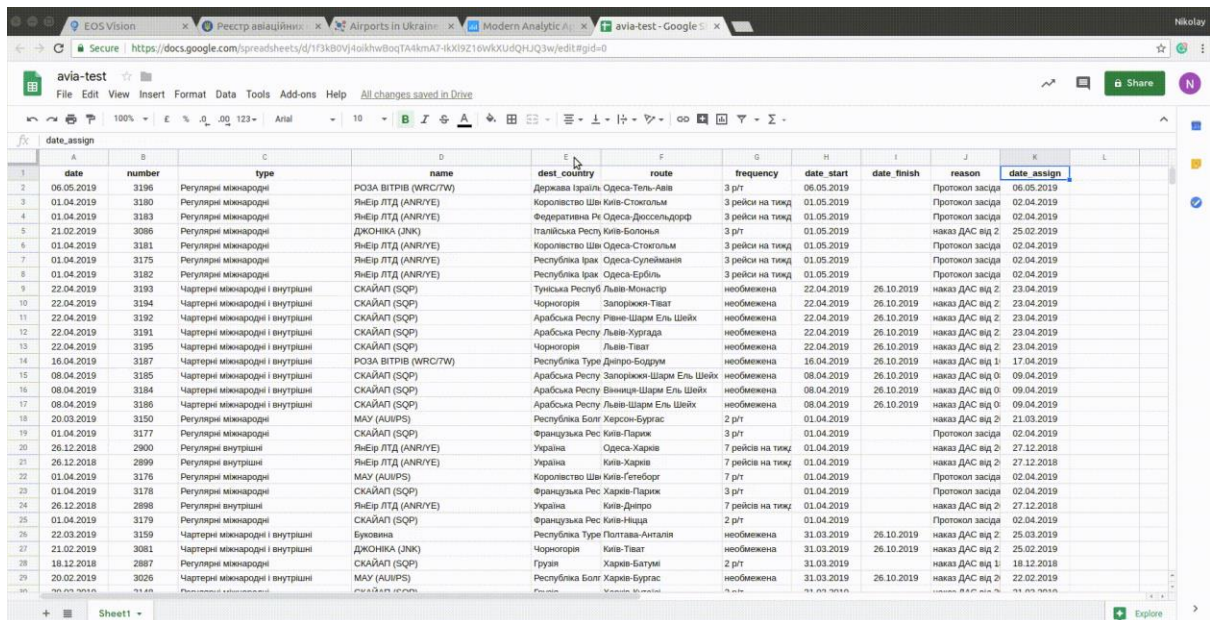
- «number» (або «building»);
- «street» (або «road»);
- «city» (або «town», або «village»);
- «neighbourhood» (або «community name», або «locality»);
- «county» (або «municipality», або «district»);
- «state» (або «region», або «province»);
- «country»;
- «postalcode» (або «postcode»).

Процес перетворення адрес в реальні точки на мапі називають геокодуванням. Сервіси, що підтримують геокодування зазвичай «розуміють» згадані назви колонок і легко перетворюють такі дані у вигляд точок на мапі.

Створення CSV

Зараз більшість сервісів підтримує роботу з даними таблиць Excel чи Open/Libre Office, а також інтегруються з хмарними сервісами Google. Однак і досі у більшості випадків така підтримка та інтеграція обмежена (наприклад, розуміння сервісом лише першої вкладки таблиці тощо). Загальноприйнятим стандартом передачі табличних даних з однієї програми в іншу є формат CSV.

Кожна з програм для роботи з табличними даними дає можливість експортувати дані в такому форматі.



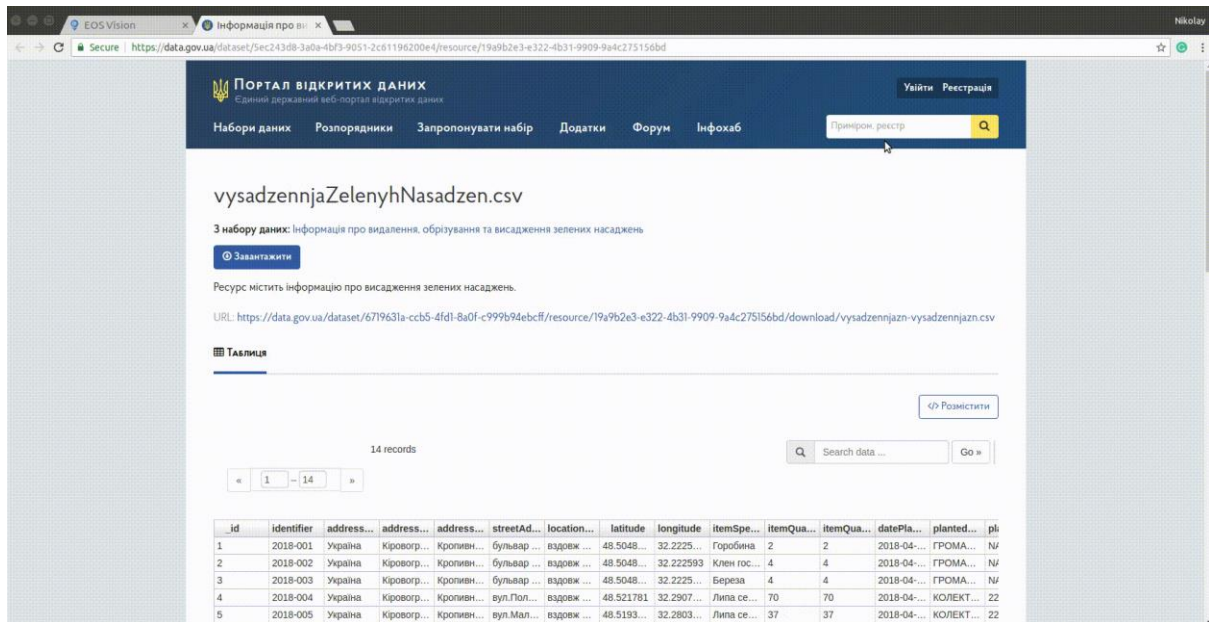
A	B	C	D	E	F	G	H	I	J	K
date	number	type	name	dest country	route	frequency	date start	date finish	reason	date assign
06.05.2019	3196	Регулярні міжнародні	РОЗА ВІТРИВ (MRC/7W)	Держава Ізраїль: Одеса-Тель-Авів	3 р/т	06.05.2019			Протокол засідд	06.05.2019
01.04.2019	3180	Регулярні міжнародні	ЯнЕір ЛТД (ANRY/E)	Королівство Шве Кіле-Стокгольм	3 рейси на тижд	01.05.2019			Протокол засідд	02.04.2019
01.04.2019	3183	Регулярні міжнародні	ЯнЕір ЛТД (ANRY/E)	Федеративна Рє Одеса-Дюссельдорф	3 рейси на тижд	01.05.2019			Протокол засідд	02.04.2019
21.02.2019	3086	Регулярні міжнародні	ДЖОНКА (JNK)	Італійська Рєспу Кіле-Болонья	3 р/т	01.05.2019			наказ ДАС вид 2	25.02.2019
01.04.2019	3181	Регулярні міжнародні	ЯнЕір ЛТД (ANRY/E)	Королівство Шве Одеса-Стокгольм	3 рейси на тижд	01.05.2019			Протокол засідд	02.04.2019
01.04.2019	3175	Регулярні міжнародні	ЯнЕір ЛТД (ANRY/E)	Рєспубліка Ірак Одеса-Сулейманья	3 рейси на тижд	01.05.2019			Протокол засідд	02.04.2019
01.04.2019	3182	Регулярні міжнародні	ЯнЕір ЛТД (ANRY/E)	Рєспубліка Ірак Одеса-Ербіль	3 рейси на тижд	01.05.2019			Протокол засідд	02.04.2019
22.04.2019	3193	Чартерні міжнародні і внутрішні	СКАЙАП (SQP)	Туніська Рєспуб Льяве-Монастір	необмежена	22.04.2019	26.10.2019	26.10.2019	наказ ДАС вид 2	23.04.2019
22.04.2019	3194	Чартерні міжнародні і внутрішні	СКАЙАП (SQP)	Чорногорія Запорожжя-Тиват	необмежена	22.04.2019	26.10.2019	26.10.2019	наказ ДАС вид 2	23.04.2019
22.04.2019	3192	Чартерні міжнародні і внутрішні	СКАЙАП (SQP)	Арабська Рєспу Раше-Шарм Ель Шейх	необмежена	22.04.2019	26.10.2019	26.10.2019	наказ ДАС вид 2	23.04.2019
22.04.2019	3191	Чартерні міжнародні і внутрішні	СКАЙАП (SQP)	Арабська Рєспу Льяве-Хургада	необмежена	22.04.2019	26.10.2019	26.10.2019	наказ ДАС вид 2	23.04.2019
22.04.2019	3195	Чартерні міжнародні і внутрішні	СКАЙАП (SQP)	Чорногорія Льяве-Тиват	необмежена	22.04.2019	26.10.2019	26.10.2019	наказ ДАС вид 2	23.04.2019
16.04.2019	3187	Чартерні міжнародні і внутрішні	РОЗА ВІТРИВ (MRC/7W)	Рєспубліка Турє Дніпро-Буддум	необмежена	16.04.2019	26.10.2019	26.10.2019	наказ ДАС вид 1	17.04.2019
08.04.2019	3185	Чартерні міжнародні і внутрішні	СКАЙАП (SQP)	Арабська Рєспу Запорожжя-Шарм Ель Шейх	необмежена	08.04.2019	26.10.2019	26.10.2019	наказ ДАС вид 0	09.04.2019
08.04.2019	3184	Чартерні міжнародні і внутрішні	СКАЙАП (SQP)	Арабська Рєспу Вінніца-Шарм Ель Шейх	необмежена	08.04.2019	26.10.2019	26.10.2019	наказ ДАС вид 0	09.04.2019
08.04.2019	3186	Чартерні міжнародні і внутрішні	СКАЙАП (SQP)	Арабська Рєспу Льяве-Шарм Ель Шейх	необмежена	08.04.2019	26.10.2019	26.10.2019	наказ ДАС вид 0	09.04.2019
20.03.2019	3150	Регулярні міжнародні	МАУ (AUI/PS)	Рєспубліка Болг Харєв-Бургас	2 р/т	01.04.2019			наказ ДАС вид 2	21.03.2019
01.04.2019	3177	Регулярні міжнародні	СКАЙАП (SQP)	Французька Рєспу Кіле-Париж	3 р/т	01.04.2019			Протокол засідд	02.04.2019
26.12.2018	2900	Регулярні внутрішні	ЯнЕір ЛТД (ANRY/E)	Україна Одеса-Харєв	7 рейсів на тижд	01.04.2019			наказ ДАС вид 2	27.12.2018
26.12.2018	2899	Регулярні внутрішні	ЯнЕір ЛТД (ANRY/E)	Україна Кіле-Харєв	7 рейсів на тижд	01.04.2019			наказ ДАС вид 2	27.12.2018
01.04.2019	3176	Регулярні міжнародні	МАУ (AUI/PS)	Королівство Шве Кіле-Гетеборг	7 р/т	01.04.2019			Протокол засідд	02.04.2019
01.04.2019	3178	Регулярні міжнародні	СКАЙАП (SQP)	Французька Рєспу Харєв-Париж	3 р/т	01.04.2019			Протокол засідд	02.04.2019
26.12.2018	2898	Регулярні внутрішні	ЯнЕір ЛТД (ANRY/E)	Україна Кіле-Дніпро	7 рейсів на тижд	01.04.2019			наказ ДАС вид 2	27.12.2018
01.04.2019	3179	Регулярні міжнародні	СКАЙАП (SQP)	Французька Рєспу Кіле-Ніцца	2 р/т	01.04.2019			Протокол засідд	02.04.2019
22.03.2019	3159	Чартерні міжнародні і внутрішні	Буковина	Рєспубліка Турє Платтава-Анталія	необмежена	31.03.2019	26.10.2019	26.10.2019	наказ ДАС вид 2	25.03.2019
21.02.2019	3081	Чартерні міжнародні і внутрішні	ДЖОНКА (JNK)	Чорногорія Кіле-Тиват	необмежена	31.03.2019	26.10.2019	26.10.2019	наказ ДАС вид 2	25.02.2019
18.12.2018	2887	Регулярні міжнародні	СКАЙАП (SQP)	Грузія Харєв-Батумі	2 р/т	31.03.2019			наказ ДАС вид 1	18.12.2018
20.02.2019	3026	Чартерні міжнародні і внутрішні	МАУ (AUI/PS)	Рєспубліка Болг Харєв-Бургас	необмежена	31.03.2019	26.10.2019	26.10.2019	наказ ДАС вид 2	22.02.2019

[Експорт даних в CSV з сервісу Google Spreadsheets](#)

[Дивитися відеоінструкцію](#)

Завантаження в інші ресурси

Приклади того, як уже оброблені дані завантажуються в різні ресурси

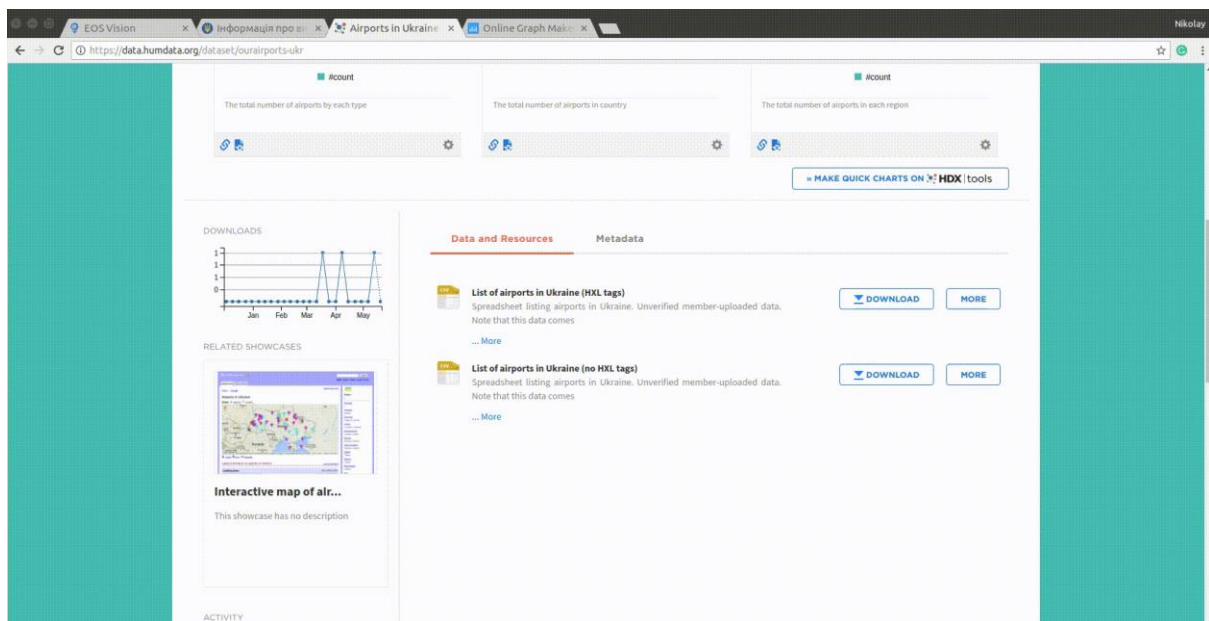


The screenshot shows the 'ПОРТАЛ ВІДКРИТИХ ДАНИХ' (Open Data Portal) interface. The main content area displays the dataset 'vysadzennjaZelenyhNasadzen.csv'. Below the dataset name, there is a 'Завантажити' (Download) button and a URL for the dataset. A 'Таблиця' (Table) section is visible, showing a table with 14 records. The table columns include 'id', 'identifier', 'address...', 'streetAd...', 'location...', 'latitude', 'longitude', 'itemSpe...', 'itemQua...', 'itemQua...', 'datePla...', and 'planted...'. The first few rows of the table are visible, showing data for tree planting in Ukraine.

id	identifier	address...	streetAd...	location...	latitude	longitude	itemSpe...	itemQua...	itemQua...	datePla...	planted...
1	2018-001	Україна	Кривоо...	Кропилен... бульвар ...	48.5048...	32.2225...	Горобина	2	2	2018-04-...	ГРОМА...
2	2018-002	Україна	Кривоо...	Кропилен... вулиця ...	48.5048...	32.222593	Клен гос...	4	4	2018-04-...	ГРОМА...
3	2018-003	Україна	Кривоо...	Кропилен... бульвар ...	48.5048...	32.2225...	Береза	4	4	2018-04-...	ГРОМА...
4	2018-004	Україна	Кривоо...	Кропилен... вул.Полт...	48.521781	32.2907...	Липа се...	70	70	2018-04-...	КОЛЕКТ...
5	2018-005	Україна	Кривоо...	Кропилен... вул.Мал...	48.5193...	32.2803...	Липа се...	37	37	2018-04-...	КОЛЕКТ...

Робота з просторовими даними у EOS Vision

[Дивитися відеоінструкцію](#)



The screenshot shows the 'Airports in Ukraine' dashboard on HumData.org. The dashboard features three charts at the top: 'The total number of airports by each type', 'The total number of airports in country', and 'The total number of airports in each region'. Below the charts, there is a 'Data and Resources' section with two entries: 'List of airports in Ukraine (HXL tags)' and 'List of airports in Ukraine (no HXL tags)'. Each entry has a 'DOWNLOAD' button and a 'MORE' button. The dashboard also includes a 'RELATED SHOWCASES' section with an 'Interactive map of air...' and an 'ACTIVITY' section.

Створення графіків в plotly

[Дивитися відеоінструкцію](#)