U-LEAD WITH EUROPE

МЕТОДОЛОГІЯ ТИПІЗАЦІЇ РЕГІОНІВ НА ОСНОВІ КЛАСТЕРНОГО АНАЛІЗУ

(апробована під час підготовки пропозицій до ДСРР2013)

Olena NYZHNYK

Short-term expert

May 2018

U-LEAD WITH EUROPE

КЛЮЧОВІ ПИТАННЯ АНАЛІЗУ РЕГІОНІВ

1. ТЕМАТИКА РЕГІОНАЛЬНОГО РОЗВИТКУ:

- Економіка
- Суспільство та навколишнє середовище
- Фактори конкурентоспроможності

2. ПРИНЦИПИ ВІДБОРУ ПОКАЗНИКІВ

- **Актуальність** – відображення різних аспектів та особливостей регіонального розвитку
- **Доступність** – на регіональному рівні (довгі динамічні ряди)
- **Надійність** – бажано не включати дані, які можуть викривлювати реальний стан справ

(наприклад, окремі показники Донецької та Луганської областей)

- **Баланс** – має бути достатньо індикаторів для кожного напрямку розвитку та баланс кількості індикаторів між кожним напрямком аналізу

Майже 100 показників

U-LEAD WITH EUROPE

ЗАГАЛЬНІ СТАТИСТИЧНІ ПІДХОДИ ТА ПОКАЗНИКИ ДЛЯ АНАЛІЗУ РЕГІОНАЛЬНИХ ДИСПРОПОРЦІЙ

- **Представлення всіх регіональних значень у відсотках від середнього по країні** і розгляд кількості регіонів вище і нижче деяких загальних порогових значень (**>125%** від середнього та **≤75%** від середнього)
- **Діапазон**, тобто **Максимальне** і **Мінімальне** значення відповідних показників, а також співвідношення між максимальним і мінімальним значеннями (**відношення макс / мін**)
- **Коефіцієнт варіації (CoV)** - показує **«загальне» відхилення від середніх значень**, на відміну від «крайніх» відхилень, зазначених максимальними і мінімальними значеннями

$$CoV = \frac{\sigma}{\bar{a}} .100$$

$$\sigma = \sqrt{\frac{\sum (a_j - \bar{a})^2}{n}}$$

CoV - коэфіцієнт варіації
σ - стандартне відхилення

a - значення показника регіону j
$\bar{a}$ - середнє значення показника по країні
n - кількість регіонів

U-LEAD WITH EUROPE

ЗАГАЛЬНІ СТАТИСТИЧНІ ПІДХОДИ ТА ПОКАЗНИКИ ДЛЯ АНАЛІЗУ РЕГІОНАЛЬНИХ ДИСПРОПОРЦІЙ

- **Динаміка змін за період у %** - **базовий період в більшості випадків (в залежності від наявності даних) - 2002-2007 рр.**
- **Ранги регіонів** - **як за 2002 і 2007 роки, так і зміна відносного положення**
- **Регіональні частки** - **відображають вагу регіонів всередині країни за конкретними показниками; у багатьох випадках були розраховані частки верхніх «n - max» і нижніх «n -min» регіонів, щоб підкреслити конкретні відмінності**
- **Були обчислені 2 варіанти: з Києвом і без Києва.**
- **Групування (класифікація) регіонів за окремими показниками за 4 групами. На основі стандартного відхилення регіональних значень від середнього по країні**

Very good	Good	Bad	Very bad
Significantly above the average	Above the average	Below the average	Significantly below the average
More than Av.+ 1/2 σ	Between Av. and Av.+ 1/2 σ	Between Av. and Av.- 1/2 σ	Less than Av. - 1/2 σ

U-LEAD WITH EUROPE

ТИПОЛОГІЯ РЕГІОНІВ

МЕТА - визначити типи регіонів, для яких будуть розроблені «шляхи розвитку» і «меню розвитку», які можуть бути основою для розробки заходів з регіонального розвитку.

ЦІЛЬ - виявлення груп регіонів (типів), які внутрішньо досить однорідні (аналогічні), але досить відрізняються від інших груп (типів), щоб виправдати диференційований підхід до втручань.

РЕКОМЕНДАЦІЯ: Для України (27 регіонів) – оптимальною кількістю є від 3 до 5 груп

Запропонована методологія охоплює як «ситуацію», так і «структуру» і «динаміку». Заснована на фактичних даних, які використовуючи об'єктивно перевіряються , що запобігає суб'єктивним апріорним судженням.

U-LEAD WITH EUROPE

КЛАСТЕРНИЙ АНАЛІЗ (ієрархічний кластерний аналіз) - ОСНОВНИЙ МЕТОД

- **Дозволяє ідентифікувати приховані групи** (типи), для яких склад та кількість груп заздалегідь не відомі.
- **Групи будуються на основі «відстаней» (різниці) між областями в багатовимірному просторі**, де кожен вимір являє змінну (індикатор), що використовується.
- **Похідні кластеру є відносно однорідними**, тобто зміна в кластері (група, тип) має бути менше, ніж зміна (відстані) між кластерами.

КЛЮЧОВІ ПЕРЕВАГИ КЛАСТЕРНОГО АНАЛІЗУ:

- здатний одночасно враховувати безліч різних вимірів і показників.
- можливість використання стандартного статистичного програмного забезпечення. В даному випадку використовувалася процедура ієрархічного кластерного аналізу пакету SPSS, яка фактично генерує початкові кластери за допомогою:
  - обчислення відстаней в багатовимірному просторі або так званої матриці близькості,
  - пропонуючи різні варіанти кількості кластерів та їх членства (тобто які регіони є членами відповідного кластера) і креслення ієрархічної дендрограми

U-LEAD WITH EUROPE

ЗМІННІ ПОКАЗНИКИ

АКТИВНІ - відносно невелика кількість ключових показників;
ПАСИВНІ – велика кількість додаткових показників, що використовуються для аналізу з метою перевірки на узгодженість.

Аналіз оснований на розгляді майже 100 індикаторів, не всі з яких були однаково інформативними, і тільки деякі з них були використані для побудови типології регіонів.

Принципи відбору активних змінних :

- має бути відносно обмежена кількість показників (занадто багато показників може «розмити» зображення, крім того багато показників корелюють, і немає ніякої необхідності включати їх всі);
- мають бути відображені різні аспекти розвитку ;
- показники повинні мати відносно високий регіональний розкид (індикатори з рівномірним розподілом тільки розмивають картину, роблячи окремі регіони дуже схожими)

U-LEAD WITH EUROPE

МЕТОДИ ІДЕНТИФІКАЦІЇ АКТИВНИХ ЗМІННИХ

- **ПОРІВНЯЛЬНИЙ АНАЛІЗ ВАРІАЦІЙ** - щоб визначити, які змінні (індикатори) більш диференційовані, ніж інші, тобто більше підходять для кластерного аналізу
- **КОРЕЛЯЦІЙНИЙ АНАЛІЗ** - для визначення змінних, які можна виключити, оскільки вони сильно корелюють з іншими показниками, що відображають той же аспект
- **КОМПОНЕНТНИЙ АНАЛІЗ** - щоб перевірити остаточний список активних змінних і особливо довести, які змінні не повинні бути виключені. Цей статистичний метод дозволяє виявити «приховані» чинники (компоненти), які пояснюють більшу частину відмінностей між регіонами

U-LEAD WITH EUROPE

ФІНАЛЬНИЙ СПИСОК АКТИВНИХ ЗМІННИХ

1. Динаміка чисельності населення, %	12. Основний капітал, тис.грн.	22. Зайнятість осіб віком 15-70 років, %
2. Щільність населення, осіб/км^2	13. Знос основних фондів, %	23. Безробіття за МОП, %
3. Вікове співвідношення населення віком 0-14/65 +, %	14. Інвестиції в основний капітал, тис.грн.	24. Вища освіта, % до 2001 р.
4. ВРП на 1 особу, тис. грн.	15. ПІІ, дол.США	25. ВДВ на 1 особу зайнятого населення, тис.грн/особу
5. Динаміка ВРП, %	16. Частка великих підприємств в обсязі реалізованої продукції, %	26. Абсолютна бідність, % (1,9 дол.США)
6. ВДВ с/г, %	17. Кількість малих підприємств на 10 тис.осіб	27. Очікувана тривалість життя при народженні, років
7. ВДВ добувна промисловість,%	18. Власні доходи місцевих бюджетів, тис.грн.	28. Смертність, тис.осіб
8. ВДВ промислове виробництво, %	19. Витрати на НДДКР, тис. грн.	29. Захворюваність на активний туберкульоз, тис.осіб
9. ВДВ будівництво,%	20. Комп'ютерне обладнання на підприємствах, од. на 100 осіб	30. % від загального забруднених стічних вод
10. ВДВ електрика, газ, вода, %	21. Кількість користувачів Інтернету (абонентів),	31. Викиди забруднюючих речовин, т/кв.км
11. ВДВ послуги, %		32. Земля під лісами, кв.км/особу

U-LEAD WITH EUROPE

ЕТАПИ ТИПОЛОГІЗАЦІЇ РЕГІОНІВ

ЕТАП 1. Визначення активних змінних (початкове групове обговорення, логічний аналіз, аналіз варіацій, кореляційний аналіз, аналіз основних компонентів)

ЕТАП 2. Запуск процедури ієрархічного кластерного аналізу на основі SPSS

ЕТАП 3. Первісна перевірка узгодженості тільки на основі активних змінних - зміна активних змінних і повторне виконання процедури

ЕТАП 4. Опис отриманих типів і перевірка узгодженості на основі всіх змінних - виявлення загальних ознак (що робить типи відмітними) в поєднанні з перевіркою узгодженості, перехресної перевіркою з докладним аналізом і частковими типологіями, в т.ч. оцінки та компоненти ІЛР і т. д.

ЕТАП 5. Експертна оцінка поправки (уточнення) на основі кроку 3 - наприклад, в основному об'єднання деяких кластерів в більші типи (якщо відмінність між кластерами була відносно невеликою), але також модифікація деяких кластерів на основі більш докладного логічного аналізу і аналізу матриці близькості

ЕТАП 6. Остаточна перевірка відповідності на основі типів засобів і відхилень

ЕТАП 7. Остаточний опис типів

U-LEAD WITH EUROPE

ЧАСТКОВІ ТИПОЛОГІЇ

- За віковою структурою
- За урбанізацією
- За ВРП в 2007 році його зростанням за 2002-2007 роки
- За типом економіки (структура ВДС по секторам)
- За капіталом та інвестиціями (як в загальному обсязі інвестицій в основний капітал, так і в обсязі ПІІ) та їх зростання в 2002-2007 рр.
- За показниками ринку праці
- За показниками якості праці (освіта)
- За показниками економічного добробуту (наявний дохід і заробітна плата і їх зростання в 2002-2007 рр., Бідність)
- За екологічними показниками

U-LEAD WITH EUROPE

ОСНОВНІ ПОКАЗНИКИ ПО РЕГІОНАХ

Значно вище середнього	Вище середнього	Нижче середнього	Значно нижче середнього

Regions	Area	Population 2009	Popul. Change 2002-2007	Migration balance 2007	Age ratio 0-14/65+ 2008	Age dependancy (0-14 + 65+/15-64) 2008	People in retirement age/ 100 employed 15-70 2007
	000 km2	000 pers.	%	pers/100 thsd.		%	%
AR of Crimea	26,1	1967,3	-2,0	179	0,91	40,73	73,3
Vinnytska	26,5	1660,0	-3,8	-91	0,83	49,26	76,5
Volynska	20,2	1036,2	-1,6	-11	1,33	47,79	69,7
Dnipropetrovska	31,9	3374,2	-3,1	25	0,81	42,21	67,8
Donetska	26,5	4500,5	-4,1	-2	0,68	42,03	70,9
Zhytomyrska	29,9	1294,2	-4,1	-133	0,92	48,63	77,2
Zakarpatska	12,8	1243,4	-0,8	-126	1,67	42,67	54,0
Zaporizka	27,3	1821,3	-3,3	-48	0,78	41,67	67,9
Ivano-Frankivska	13,9	1381,1	-1,3	-15	1,19	45,34	73,8
Kyivska	28,1	1727,8	-3,2	-42	0,87	44,18	73,6
Kirovogradska	24,6	1027,0	-5,6	-401	0,80	46,42	75,1
Luganska	26,7	2331,8	-5,0	-204	0,68	40,80	76,3
Lvivska	21,8	2552,9	-1,6	-47	1,08	43,30	68,0
Mykolaivska	24,6	1195,8	-3,2	-53	0,95	41,58	65,0
Odeska	33,3	2392,2	-2,2	473	1,00	41,35	63,3
Poltavska	28,8	1511,4	-4,3	-37	0,73	45,01	74,9
Rivnenska	20,1	1151,0	-1,2	-171	1,47	47,58	73,0
Sumska	23,8	1184,0	-5,4	-193	0,72	43,53	73,4
Ternopilska	13,8	1093,3	-2,5	-155	0,98	47,29	83,8
Kharkivska	31,4	2782,4	-2,6	144	0,76	38,68	64,0
Khersonska	28,5	1099,2	-3,8	-269	0,99	42,40	66,3

U-LEAD WITH EUROPE

ЧАСТКОВІ ТИПОЛОГІЇ

1. Київ – особливий випадок як у кількісному, так і у якісному відношенні, котрий неможливо порівняти або об’єднати з будь-яким іншим регіоном.

2. Високоіндустріалізовані та високоурбанізовані регіони з рівнем розвитку вище середнього (з 2 підтипами) – дуже компактний тип, що охоплює регіони у східній частині країни.

3. Високоурбанізовані, орієнтовані на послуги регіони з середнім рівнем розвитку (з 2 підтипами) – регіони у різних частинах країни, з більш компактною групою на півдні.

4. Малі, дещо урбанізовані, агропромислові регіони з низьким рівнем розвитку (з 3 підтипами) – відносно компактний тип, здебільшого у західній частині країни.

5. Регіони з низьким рівнем розвитку, вповільненням розвитку та з найсерйознішими проблемами розвитку – розкиданий тип.

U-LEAD WITH EUROPE

ПРОПОЗИЦІЇ

1. Методологію взяти за основу типологізації регіонів для ДСРР2027.
2. Переглянути перелік показників (активні + пасивні), визначити власний список індикаторів для аналізу.
3. Вивчити програмний пакет SPSS.
4. Розпочати аналіз.
5. Визначити типи територій для формування «меню втручань» до Плану заходів з реалізації ДСРР2027

Funded by the European Union

співпраця з німеччиною
DEUTSCHE ZUSAMMENARBEIT

Sweden Sverige

Polish aid

ESTONIA
DEVELOPMENT COOPERATION

THE MINISTRY OF REGIONAL DEVELOPMENT, CONSTRUCTION, HOUSING AND COMMUNAL SERVICES

U-LEAD

WITH EUROPE

THANK YOU!