

План-конспект лекції: Роль маркетингових досліджень у глобальній бізнес-стратегії

Навчальні цілі лекції:

- Надати коректне визначення дескриптивної аналітики та розказати про її мету в роботі аналітика.
- Пояснити, як відбувається аналітичний цикл «дані — опис — пояснення — прогноз — рекомендація» та яке місце дескриптивної аналітики в ньому.
- Розповісти, які бувають джерела даних.
- Деталізувати методи дескриптивної аналітики (агрегування, порівняння, групування, візуалізація), показати типові дашборди.
- Розглянути типові помилки інтерпретації описових даних.
- Показати вплив дескриптивної аналітики на продуктові рішення через реальні кейси.

Очікувані результати навчання:

Після лекції студент:

- коректно формулює визначення та мету дескриптивної аналітики;
- пояснює місце дескриптивної аналітики в аналітичному циклі;
- розрізняє джерела та методи обробки описових даних;
- формує структуру типових дашбордів;
- ідентифікує поширені помилки інтерпретації даних;
- наводить приклади впливу описових інсайтів на рішення команди продукту.

Структура лекції:

- **Вступ. Роль і значення дескриптивної аналітики**
 - Навіщо потрібна дескриптивна аналітика
 - Які завдання виконує дескриптивна аналітика
- **Аналітичний цикл та місце дескриптивної аналітики**
 - Gartner Analytic Ascendancy Model
 - 4 етапи використання даних
- **Інструментарій дескриптивної аналітики**
 - Джерела даних для дескриптивної аналітики
 - Внутрішні джерела
 - Зовнішні джерела
 - Ключові методи: агрегування, порівняння, групування, візуалізація даних
- **Практичне застосування та звіти у дескриптивній аналітиці**
 - Дашборди в продуктовому IT
 - Фреймворк AARRR
- **Інтерпретація та поширені помилки**

- Парадокс Сімпсона (Simpson's Paradox)
 - Кейс: гендерна упередженість в Університеті Берклі (1973)
- Систематична помилка того, хто вижив (Survivorship Bias)
 - Кейс: бомбардувальники Другої світової війни
- Correlation vs. Causation
- **Вплив на продуктові рішення (кейс-стаді)**
 - Кейс 1: оптимізація воронки конверсії в SaaS-продукті
 - Кейс 2: персоналізація контенту в медіа-сервісі (на прикладі Netflix)
- **Висновок**
- **Список джерел**

Очікувані теми попереднього опрацювання (prerequisites): Для успішного засвоєння матеріалу лекції студенти повинні мати такі попередні знання та вміння:

- **Базові статистичні поняття:** розуміння ключових метрик описової статистики, таких як середнє значення, медіана, а також вміння працювати з відсотками та пропорціями.
- **Базові знання аналізу даних:** уявлення про те, що таке дані, які є їхні типи (кількісні, якісні) та як вони можуть бути структуровані.
- **Основи візуалізації даних:** знайомство з основними типами діаграм (наприклад, лінійний графік, стовпчикова та кругова діаграми) та розуміння їхнього призначення.
- **Логічне мислення й аналітичні здібності:** здатність аналізувати базову інформацію, виявляти закономірності та робити обґрунтовані висновки.

Окрім цього, до цієї лекції студентам рекомендовано пройти **модуль 1** всеукраїнського курсу «Аналітика у продуктовому IT».

Вступ. Роль та значення дескриптивної аналітики

Навіщо потрібна дескриптивна аналітика?

Дескриптивна (описова) аналітика — це процес аналізу історичних даних для підсумовування та виявлення патернів, трендів і ключових показників. Її фундаментальна мета — надати чітку, засновану на фактах відповідь на питання «Що сталося?». Вона не прогнозує майбутнє і не шукає глибинних причин, а створює об'єктивну картину минулого, перетворюючи величезні масиви необроблених даних на зрозумілі інсайти.

Дескриптивна аналітика є фундаментом, на якому будуються **інші рівні аналізу**: діагностичний, предиктивний та прескриптивний. Якість усіх подальших, складніших аналітичних висновків безпосередньо залежить від точності та релевантності даних, отриманих на цьому початковому етапі. Принцип «сміття на вході — сміття на виході» тут

є критичним, адже помилки в описових даних неминуче призведуть до хибних прогнозів та неефективних рекомендацій.

У щоденній роботі фахівця дескриптивна аналітика виконує 3 ключові завдання:

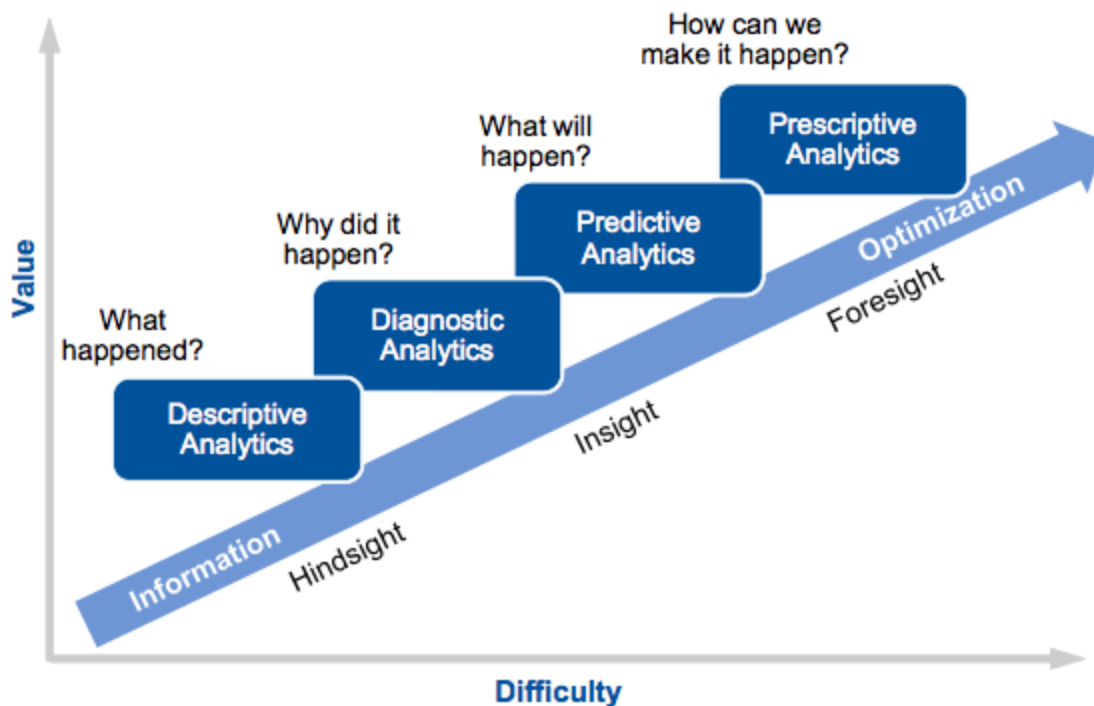
- **Моніторинг бізнес-показників.** Це безперервне відстеження ключових показників ефективності (KPI): щоденні та місячні активні користувачі (DAU/MAU), дохід (MRR/ARR), коефіцієнт конверсії, рівень утримання (Retention Rate) та відтоку клієнтів (Churn Rate).
- **Регулярна звітність.** Це створення й автоматизація щоденних, щотижневих та щомісячних звітів і дашбордів, які надають актуальну інформацію для всіх зацікавлених сторін — від маркетингової команди до топменеджменту.
- **Виявлення трендів та аномалій.** Це ідентифікація стійких патернів у поведінці користувачів (наприклад, сезонне зростання попиту на певну функцію) та несподіваних відхилень від норми (наприклад, різке падіння залученості після оновлення застосунку), які сигналізують про необхідність глибшого дослідження.

Аналітичний цикл і місце дескриптивної аналітики

Цикл аналітики: від даних до рекомендацій (Gartner Analytic Ascendancy Model)

Щоб повною мірою зрозуміти місце дескриптивної аналітики, необхідно розглянути її в контексті повного аналітичного циклу. Наприклад, можна використати **модель аналітичної зрілості Gartner (Gartner Analytic Ascendancy Model)**. Ця модель описує 4 етапи, які проходить організація на шляху до повноцінного використання даних.

Gartner, Inc — світова дослідницька і консалтингова компанія у сфері інформаційних технологій. Заснована у 1979, у США.



Source: Gartner (March 2012)

Докладний опис чотирьох етапів:

1. Етап 1: описативна аналітика (опис) — що сталося?

- **Мета:** надати об'єктивний зріз історичних даних. Це відправна точка, яка констатує факти без пояснення причин.
- **Техніки:** агрегація даних, розрахунок зведених статистик (середнє арифметичне, медіана), створення статичних звітів та інтерактивних дашбордів.
- **Приклад у роботі:** звіт показує: «минулого місяця показник утримання користувачів у застосунку на 7-й день (Day-7 Retention) склав 25%». Він зменшився на 5,3%.

2. Етап 2: діагностична аналітика (пояснення) — чому це сталося?

- **Мета:** заглибитися в дані, щоб знайти першопричини подій, виявлених на першому етапі. Це етап розслідування, який шукає взаємозв'язки та аномалії.
- **Техніки:** деталізація даних (Drill-down), аналіз кореляцій, виявлення викидів, сегментація аудиторії.
- **Приклад у роботі:** аналітик досліджує падіння утримання і виявляє, що зниження показника збіглося з випуском нової версії застосунку, яка мала критичний баг у процесі онбордингу для користувачів на платформі Android.

3. Етап 3: предиктивна аналітика (прогноз) — що може статися?

- **Мета:** використовувати історичні дані для побудови моделей, які прогнозують майбутні результати та оцінюють ймовірності різних сценаріїв.
- **Техніки:** статистичне моделювання, алгоритми машинного навчання (регресія, класифікація), прогнозування часових рядів.
- **Приклад у продукті:** модель прогнозує, що на основі поточних тенденцій відтоку, компанія очікує, що в наступному кварталі кількість активних користувачів зменшиться на 15%, якщо проблему з онбордингом не буде розв'язано.

4. Етап 4: прескриптивна аналітика (рекомендація) — що слід робити?

- **Мета:** не просто прогнозувати майбутнє, а й рекомендувати конкретні дії для досягнення бажаного результату або уникнення небажаного. Це найвищий рівень аналітичної зрілості.
- **Техніки:** оптимізаційні алгоритми, симуляції (what-if аналіз), A/B-тестування, системи рекомендацій.
- **Приклад у продукті:** аналітик моделює різні варіанти дій і рекомендує запустити цільову email-кампанію для сегмента користувачів Android, які зіткнулися з багом, пропонуючи їм бонус для повернення в додаток. Симуляція показує, що це може підвищити утримання на 5% і є більш ефективним, ніж загальне повідомлення для всіх користувачів.

Інструментарій дескриптивної аналітики

Джерела даних для дескриптивної аналітики:

- **Внутрішні джерела:** це дані, що генеруються всередині компанії та її продуктів. Вони є основним джерелом для продуктової аналітики.
 - Дані про поведінку користувачів: логи з мобільних застосунків і вебсайтів, які фіксують кожну дію користувача — кліки, перегляди екранів, час сесії, використання конкретних функцій, послідовність дій. Це найцінніші дані для розуміння взаємодії з продуктом.
 - Транзакційні дані: інформація з систем управління взаємовідносинами з клієнтами (CRM) та планування ресурсів підприємства (ERP). Сюди входять дані про покупки, оформлення підписок, історію платежів, скасування замовлень.
 - Дані з систем підтримки: тексти звернень до служби підтримки, відгуки, скарги, оцінки задоволеності обслуговуванням. Ці дані допомагають виявити проблеми, з якими зіштовхуються користувачі.
 - Фінансові дані: звіти про доходи, витрати, прибутковість продукту, які дозволяють оцінити його економічну ефективність.
- **Зовнішні джерела:** це дані, що збираються з-поза меж компанії для отримання ширшого ринкового контексту та доповнення внутрішніх даних.

- Дані з соціальних мереж: аналіз лайків, коментарів, репостів і згадок бренду для оцінювання настроїв аудиторії (sentiment analysis) та ефективності маркетингових кампаній.
- Ринкові дослідження: звіти аналітичних агентств, дані про частку ринку, аналіз конкурентів, їхніх продуктів і цінової політики.
- Опитування користувачів: прямий збір зворотного зв'язку від цільової аудиторії через анкети та інтерв'ю для вимірювання задоволеності (CSAT, NPS), виявлення потреб і тестування нових ідей.
- Дані з відкритих API: інформація від третіх сторін, яка може впливати на поведінку користувачів, наприклад, дані про погоду (для туристичних застосунків), курси валют (для фінтех-продуктів) або демографічна статистика.

Необроблені дані рідко бувають готовими до аналізу. Процес **підготовки даних (Data Preparation)** є важливим етапом. Він охоплює:

1. **Збір та агрегацію (Data Aggregation):** процес збору даних з різноманітних джерел (бази даних, CRM, лог-файли) в єдине централізоване сховище, наприклад, Data Warehouse.
2. **Очищення (Data Cleaning):** виявлення та виправлення помилок у даних, обробка пропущених значень, видалення дублікатів, корекція невірних записів і стандартизація форматів.
3. **Трансформацію (Data Transformation):** перетворення даних у зручний для аналізу формат. Це може включати розрахунок нових похідних показників (наприклад, віку користувача з його дати народження) або зміну структури даних для подальшого моделювання.

Ключові методи: агрегування, порівняння, групування, візуалізація

Після підготовки даних аналітики застосовують набір основних методів для виявлення значущих інсайтів.

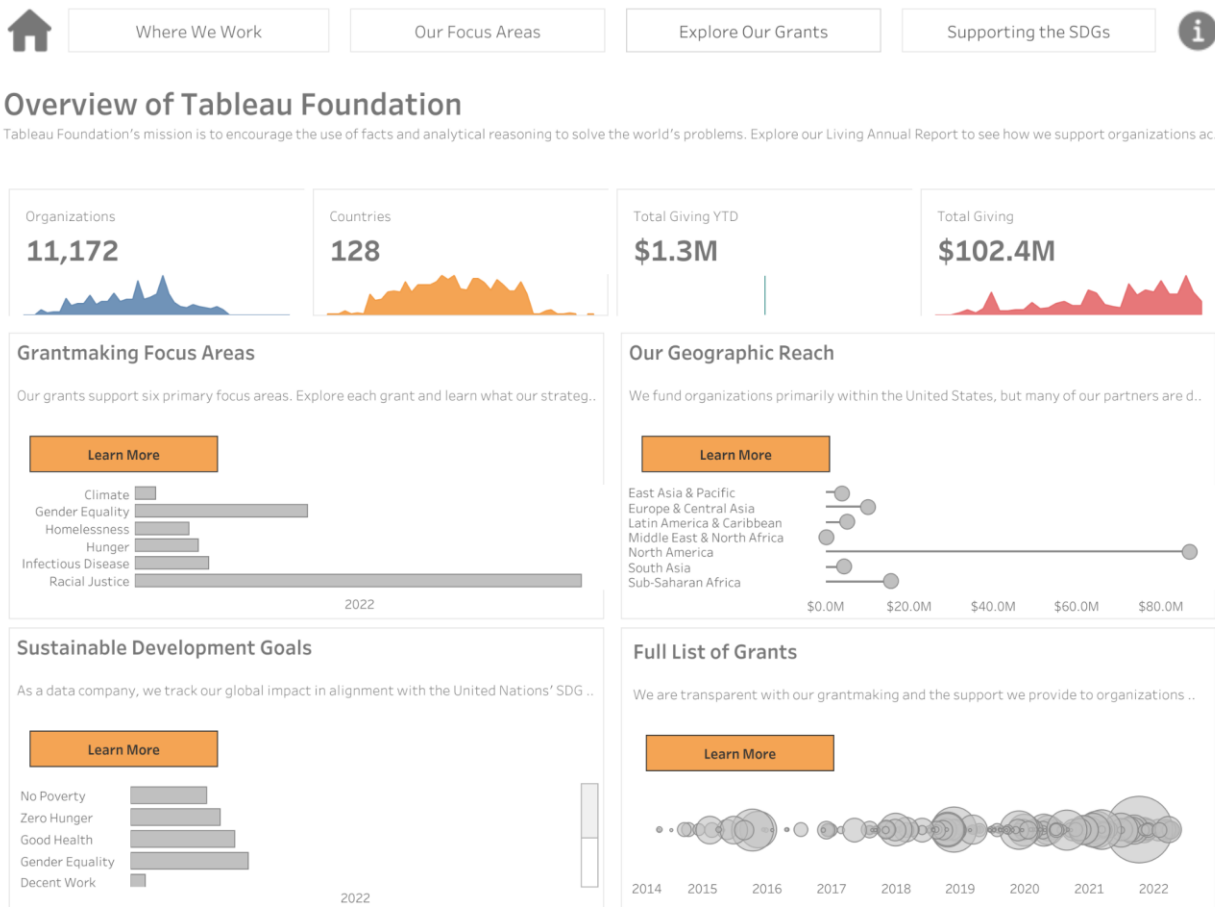
- **Агрегування:** це процес зведення деталізованих даних до підсумкових, узагальнених показників.
 - Міри центральної тенденції: **середнє**, **медіана** та **мода** використовуються для визначення типового або центрального значення в наборі даних.
 - Міри розсіювання (дисперсії): **діапазон**, **стандартне відхилення** та **квартилі** допомагають оцінити, наскільки дані мінливі та розкидані навколо центрального значення.

- **Групування та сегментація:** це розподіл гетерогенної сукупності даних на однорідні групи (сегменти) для глибшого аналізу. Сегментація дозволяє виявити патерни, які залишаються невидимими при аналізі загальних, агрегованих даних.
 - *Приклади сегментації в продукті:*
 - За демографією: вік, стать, країна.
 - За поведінкою: power users (найбільш активні), користувачі, що ризикують піти (зменшили активність), нові користувачі.
 - За платформою: iOS, Android, Web.
 - За джерелом залучення: органічний пошук, платна реклама, соціальні мережі.
 - За тарифним планом: Free, Pro, Enterprise.
- **Порівняння:** дані набувають сенсу лише в контексті. Порівняння є ключовим методом для створення цього контексту.
 - Динамічне порівняння: аналіз показників у часі (день до дня, місяць до місяця, рік до року) для виявлення трендів, сезонності та темпів зростання.
 - Бенчмаркінг: порівняння власних показників із середніми по галузі, показниками прямих конкурентів або внутрішніми цілями (KPI) для оцінки ефективності своєї роботи.
- **Візуалізація:** це представлення даних у графічному форматі (графіки, діаграми, карти), що значно полегшує їх сприйняття, інтерпретацію та комунікацію результатів аналізу нетехнічним стейкхолдерам. Вибір правильного типу візуалізації є критично важливим для ефективного передачі інформації.
 - **Лінійний графік (line chart):** ідеальний для відстеження динаміки одного або кількох показників у часі (наприклад, динаміка DAU протягом місяця).
 - **Стовпчикова діаграма (bar chart):** використовується для порівняння значень між різними дискретними категоріями (наприклад, дохід по країнах або ефективність різних маркетингових каналів).
 - **Гістограма (histogram):** призначена для аналізу розподілу неперервних числових даних. Вона показує, як часто значення потрапляють у певні діапазони. Наприклад, розподіл користувачів за кількістю покупок.
 - **Секторна діаграма (pie chart):** показує пропорції або частки від цілого. Її слід використовувати обережно, лише коли категорій небагато (зазвичай не більше 5–6), інакше вона стає нечитабельною.
 - **Теплова карта (heat map):** використовує колір для візуалізації щільності даних у матриці. Часто застосовується для аналізу активності користувачів на вебсторінці (куди найчастіше клікають).
 - **Коробковий графік (box plot):** надає стислий огляд розподілу даних, показуючи медіану, квартилі та викиди. Дуже корисний для порівняння розподілів між кількома сегментами.

Практичне застосування та звіти у дескриптивній аналітиці

Дашборди в продуктовому IT

Дашборд — це інструмент візуалізації, який в інтерактивному режимі відображає набір ключових метрик і показників ефективності (KPI), дозволяючи швидко та комплексно оцінити стан продукту, бізнес-процесу чи компанії в цілому.

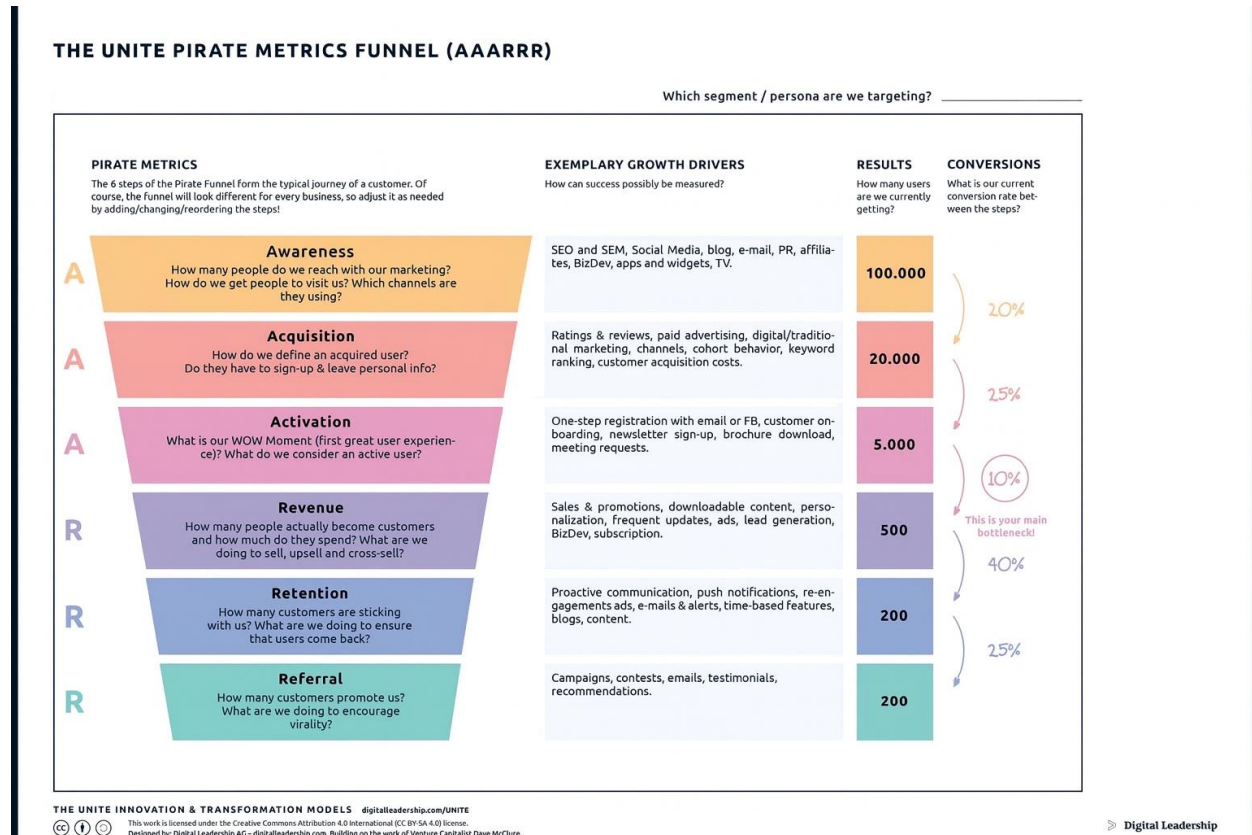


Приклад дашборду від Tableau.

На відміну від статичних звітів, які фіксують стан на певний момент часу (наприклад, квартальний фінансовий звіт), дашборди зазвичай оновлюються в режимі, близькому до реального часу, і надають можливість для деталізації даних (drill-down).

Для продуктового IT одним із ефективних підходів до структурування дашборду є використання **фреймворку AARRR**, також відомого як «**Піратська воронка**». Ця модель, запропонована інвестором Дейвом МакКлюром, розглядає життєвий цикл користувача як послідовну воронку з п'яти етапів, кожен з яких має свої ключові метрики. Такий підхід

створює цілісну розповідь про шлях користувача, дозволяючи швидко діагностувати, на якому етапі виникають проблеми.



Докладний розбір продуктового дашборду на основі фреймворку AARRR:

1. Acquisition (залучення) — як користувачі нас знаходять?

- **Стратегічне питання:** які маркетингові канали приводять найбільшу кількість якісних користувачів за оптимальною ціною?
- **Ключові метрики на дашборді, які можна використати**
 - **Кількість нових користувачів/реєстрацій:** відображається в динаміці (за днями/тижнями) та з розбивкою за каналами (Organic Search, Paid Ads, Social Media, Direct).
 - **Вартість залучення клієнта (Customer Acquisition Cost, CAC):** розраховується для кожного каналу, щоб оцінити його економічну ефективність.
 - **Коефіцієнт конверсії лендінгу:** відсоток відвідувачів сайту, які виконали цільову дію (наприклад, зареєструвалися).

2. Activation (активація) — чи відчують користувачі першу цінність?

- **Стратегічне питання:** чи виконують нові користувачі ключові дії, які корелюють з довгостроковим використанням продукту та переходом до оплати?
- **Ключові метрики на дашборді:**
 - **Activation Rate:** відсоток нових користувачів, які виконали заздалегідь визначену ключову дію.
 - **Time-to-Value (TTV):** середній час, який проходить від моменту реєстрації до моменту активації. Чим він менший, тим краще.
 - **Відсоток завершення онбордингу:** якщо в продукті є покроковий процес знайомства, ця метрика показує, скільки користувачів його проходять до кінця.

3. Retention (утримання) — чи повертаються користувачі?

- **Стратегічне запитання:** чи є наш продукт достатньо цінним, щоб користувачі поверталися до нього знову і знову?
- **Ключові метрики на дашборді:**
 - **Retention Rate:** зазвичай виконується у вигляді когортного аналізу, який показує, який відсоток користувачів, що зареєструвалися в певний період (когорти), повертається в продукт на 1-й, 7-й, 30-й день тощо.
 - **Churn Rate (коефіцієнт відтоку):** відсоток користувачів, які припинили використовувати продукт за певний період. Це зворотна до утримання метрика.
 - **Співвідношення DAU/MAU:** частка щомісячних активних користувачів, які взаємодіють з продуктом щодня. Високий показник (наприклад, > 40%) свідчить про високу залученість у продукт.

4. Referral (реферали) — чи рекомендують нас користувачі?

- **Стратегічне питання:** чи настільки задоволені наші користувачі, що готові стати амбасадорами нашого бренду і залучати нових клієнтів?
- **Ключові метрики на дашборді:**
 - **Net Promoter Score (NPS):** індекс споживчої лояльності, що вимірюється через опитування. Показує співвідношення «промоутерів», «нейтралів» і «критиків».
 - **Кількість надісланих запрошень:** пряма метрика використання реферальної функціональності.
 - **Віральний коефіцієнт (K-factor):** показник, що розраховується як кількість запрошень, надісланих одним користувачем * коефіцієнт конверсії запрошення. Якщо $K > 1$, продукт має потенціал до експоненційного зростання.

5. Revenue (дохід) — як ми заробляємо гроші?

- **Стратегічне питання:** чи готові користувачі платити за цінність, яку ми надаємо, і наскільки прибутковою є наша бізнес-модель?
- **Ключові метрики на дашборді:**
 - **Місячний регулярний дохід (Monthly Recurring Revenue, MRR):** ключова метрика для підпискових бізнесів.
 - **Довічна цінність клієнта (Lifetime Value, LTV):** прогнозований загальний дохід, який компанія отримає від одного клієнта за весь час його взаємодії з продуктом.
 - **Середній дохід на користувача (Average Revenue Per User, ARPU):** загальний дохід, поділений на кількість користувачів.
 - **Співвідношення LTV/CAC:** один з найважливіших показників здоров'я бізнесу. Вважається, що для стійкої моделі це співвідношення має бути не менше 3:1.

Інтерпретація та поширені помилки

Дескриптивна аналітика є потужним інструментом, але варто знати, як коректно інтерпретувати дані. Неправильна інтерпретація може призвести до хибних бізнес-рішень з серйозними фінансовими наслідками.

Парадокс Сімпсона (Simpson's Paradox)

Парадокс Сімпсона — це статистичний феномен, за якого тренд, що спостерігається в кількох окремих групах даних, зникає або навіть змінюється на протилежний при їх об'єднанні (агрегації). Причиною цього парадоксу є наявність **прихованої змінної (confounding variable)**, яка нерівномірно розподілена між групами і суттєво впливає на результат.

Кейс: гендерна упередженість в Університеті Берклі (1973)

1. **Агреговані дані:** при аналізі загальних даних про вступ до аспірантури виявилося, що відсоток вступу для чоловіків (44%) був значно вищим, ніж для жінок (35%). Це створювало враження явної гендерної дискримінації.

	All		Men		Women	
	Applicants	Admitted	Applicants	Admitted	Applicants	Admitted
Total	12,763	41%	8,442	44%	4,321	35%

2. **Дезагреговані дані:** коли дослідники проаналізували дані по кожному факультету окремо, картина кардинально змінилася. На більшості факультетів відсоток вступу для жінок був таким самим або навіть вищим, ніж для чоловіків.

Department	All		Men		Women	
	Applicants	Admitted	Applicants	Admitted	Applicants	Admitted
A	933	64%	825	62%	108	82%
B	585	63%	560	63%	25	68%
C	918	35%	325	37%	593	34%
D	792	34%	417	33%	375	35%
E	584	25%	191	28%	393	24%
F	714	6%	373	6%	341	7%
Total	4526	39%	2691	45%	1835	30%

Legend:

- greater percentage of successful applicants than the other gender
- greater number of applicants than the other gender

bold - the two 'most applied for' departments for each gender

3. **Прихована змінна: факультет, на який подавали документи.** Виявилося, що жінки частіше обирали більш конкурентні факультети (наприклад, англійську філологію), де загальний відсоток вступу був низьким для обох статей. Чоловіки ж частіше подавали документи на менш конкурентні технічні факультети з високим відсотком вступу. Агрегація даних приховувала цей важливий фактор, створюючи ілюзію упередженості.

Систематична помилка того, хто вижив (Survivorship Bias)

Помилка того, хто вижив — це логічна хиба, що полягає в тому, щоб робити висновки на основі даних, які успішно пройшли певний процес відбору, ігноруючи при цьому дані, які цей відбір не пройшли. Це призводить до надмірно оптимістичних або просто неправильних висновків.

Класичний приклад: бомбардувальники Другої світової війни

1. **Проблема:** військове командування США хотіло зрозуміти, де потрібно зміцнити броню на бомбардувальниках, щоб підвищити їхню витривалість.
2. **Аналіз «тих, хто вижив»:** інженери аналізували літаки, які поверталися з бойових вильотів, і наносили на схему місця пробойн. Вони помітили, що найбільше пошкоджень було на крилах, фюзеляжі та хвостовій частині, тоді як двигуни та кабіна пілота залишалися практично неущкодженими.
3. **Хибний висновок:** початкова пропозиція полягала в тому, щоб зміцнювати ті частини, де було найбільше пробойн.
4. **Правильний висновок (математик Абрахам Вальд):** Вальд вказав на критичну помилку в логіці. Дані були неповними, оскільки аналізувалися лише літаки, що *повернулися*. Він стверджував, що зміцнювати потрібно якраз ті місця, де пробойн *не було* — двигуни та кабіну. Літаки, уражені в цих критичних зонах, просто не поверталися на базу і, відповідно, не потрапляли в аналіз.

Correlation vs. Causation

Це, мабуть, найпоширеніша пастка в інтерпретації даних. Дескриптивна аналітика чудово виявляє **кореляції** — статистичні зв'язки між двома або більше змінними. Однак сам факт наявності кореляції **не означає** наявності причинно-наслідкового зв'язку.



- **Приклад у продукті:** аналітик бачить на дашборді, що користувачі, які використовують нову функцію, мають на 30% вищий показник утримання (Retention Rate), ніж ті, хто її не використовує.
- **Хибний висновок:** «Функція X змушує користувачів залишатися в продукті, тож потрібно змусити всіх її використовувати».
- **Можливі альтернативні пояснення:**
 1. **Зворотна причинність:** можливо, не функція підвищує утримання, а найбільш лояльні та залучені користувачі (які й так би залишилися) просто схильні пробувати всі нові функції.
 2. **Наявність третьої змінної:** можливо, і використання функції, і високе утримання спричинені третім фактором (наприклад, ці користувачі пройшли якісніший онбординг або належать до певного професійного сегмента).

Дескриптивна аналітика, за своїм визначенням, може лише **констатувати співіснування подій**. Щоб довести причинність, необхідний перехід до діагностичного та експериментального аналізу (наприклад, проведення A/B-тесту). Тому основна роль описового звіту — не надавати готові **відповіді**, а генерувати добре сформульовані **гіпотези** для подальшої перевірки.

Вплив на продуктові рішення (кейс-стаді)

Розглянемо кілька практичних кейсів, як використовують описову аналітику в роботі.

Кейс 1: оптимізація воронки конверсії в SaaS-продукті

Проблема: Продуктова команда помічає, що SaaS-сервіс має високий трафік і велику кількість реєстрацій, але дуже низький відсоток користувачів переходить на платний тарифний план. Загальний показник конверсії є незадовільним.

Застосування описової аналітики:

1. **Побудова воронки конверсії:** аналітики використовують інструменти для візуалізації ключових етапів шляху користувача, від першого контакту до оплати.
2. **Аналіз відвалу (Drop-off):** дашборд з воронкою чітко показує, що на етапі «Створення першого проєкту» відпадає 80% користувачів, які зареєструвалися. Це і є головне вузьке місце (bottleneck) у воронці.
3. **Перехід до діагностики:** щоб зрозуміти, *чому* користувачі не створюють проєкт, аналітики використовують інструменти для аналізу поведінки, такі як записи сесій (session recordings) та теплові карти (heatmaps) для цього конкретного екрана. Аналіз показує, що користувачі губляться в складному інтерфейсі, не розуміють, які поля заповнювати, і залишають сторінку.
4. **Рішення та вплив:** на основі цих інсайтів продуктова команда повністю переробляє та спрощує інтерфейс створення проєкту, додаючи підказки та покрокову інструкцію. Через місяць після оновлення повторний описовий аналіз воронки показує, що конверсія на цьому етапі зросла з 20% до 45%. Це безпосередньо призвело до збільшення кількості користувачів, що доходять до етапу оплати, та зростання місячного регулярного доходу (MRR).

Кейс 2: персоналізація контенту в медіа-сервісі (на прикладі Netflix)

Проблема: як підвищити залученість та утримання користувачів у сервісі, що пропонує тисячі фільмів і серіалів, і допомогти їм знайти релевантний контент серед великої кількості інформації?

Застосування описової аналітики:

1. **Масштабний збір даних:** Netflix збирає величезні обсяги історичних даних про перегляди: що саме дивляться користувачі, коли, як довго, на якому пристрої, які жанри, яких акторів, у якій послідовності.
2. **Кластеризація:** за допомогою методів описової аналітики (зокрема, кластерного аналізу) Netflix групує користувачів у тисячі мікросегментів (taste

clusters). Ці кластери є набагато більш точними та гранулярними, ніж прості жанри. Наприклад, замість жанру «фантастика», кластер може бути «похмурі науково-фантастичні фільми 90-х з сильним жіночим персонажем і філософським підтекстом».

3. **Рішення та вплив:** на основі цих описуваних кластерів система рекомендацій Netflix пропонує користувачам контент, який історично був популярним серед інших членів їхнього taste cluster. Це не так предиктивна аналітика («ми прогнозуємо, що вам сподобається»), як описувана («ви належите до групи, якій історично подобався цей контент»). Такий підхід до персоналізації різко підвищує залученість, час, проведений у сервісі, і є одним з ключових факторів надзвичайно високого рівня утримання клієнтів Netflix.

Описувана аналітика є незамінною частиною у роботі будь-якої data-driven продуктової компанії. Вона виконує критичну функцію перетворення сирих даних на структуровану, зрозумілу картину минулого, відповідаючи на фундаментальне питання «Що сталося?».

Список додаткових джерел для самостійного опрацювання

1. [What Is Descriptive Analytics?](#)
2. [Sex Bias in Graduate Admissions: Data from Berkeley](#)
3. [How Taste Clusters Influence the Books You Watch on Netflix](#)
4. [A Deep Dive Into Recommendation Algorithms With Netflix](#)
5. [Descriptive Analytics: Understanding the Past to Inform the Future](#)
6. [Data Camp. The Four Types of Analytics](#)
7. [Tableau. What is Descriptive Analytics?](#)
8. [Harvard Business School. 4 Examples of Business Analytics in Action](#)
9. [Lean Analytics: Use Data to Build a Better Startup Faster. Алістер Кролл, Беньямін Йосковіц](#)
10. [Dave McClure. Startup Metrics for Pirates](#)
11. [Harvard Business School. 4 Types of Data Analytics to Improve Decision-Making](#)
12. [Khan Academy](#)
13. [Курс зі статистики від Гарвардського університету](#)
14. [Ютуб-канал з анімованими відео про математику](#)
15. [Курс про теорію ймовірності від Массачусетського університету технологій](#)
16. [Книга для самостійного вивчення статистики з вправами](#)