Пәні: Деректерді интеллектуалды талдау және визуализациялау

Деректерді талдау және визуализациялау негіздері

Аналитикалық ойлау, сапалы дерек және түсінікті визуалды коммуникация

Дәріс 6

Шаяхметова А.С., PhD, қауымдастырылған профессор
Жасанды интеллект және Big Data кафедрасы

2026-2027 оқу жылы

Мақсаты

- Деректерді талдаудың негізгі кезеңдерін жүйелі түрде түсіндіру.
- Дерек типтері, сапасы және сипаттамалық статистиканы ажырату.
- Визуализация түрін дерек пен мақсатқа сәйкес таңдауды көрсету.
- Нәтижені кесте, график және дашборд арқылы дұрыс ұсыну қағидаларын қалыптастыру.

Бұл тақырып деректі жинаудан бастап оны түсінікті визуалды қорытындыға айналдыруға дейінгі толық циклді қамтиды.

Жоспар

№	Жоспар пункті	Негізгі мазмұн
	Деректерді талдау процесі	мақсат, жинау, тазалау, зерттеу, қорытынды
	Статистикалық сипаттамалар және сапа	орташа, медиана, таралу, корреляция, аномалия
	Визуализация негіздері	график түрлері, дизайн қағидалары, қателер
	Құралдар мен практикалық қолдану	Excel, Python, Power BI, дашборд және есеп

1. Деректерді талдау процесі

Бұл бөлімде деректерді талдау мақсат қоюдан басталып, дайын қорытынды мен ұсынысқа дейінгі қадамдық үдеріс ретінде түсіндіріледі.



Аналитикалық цикл



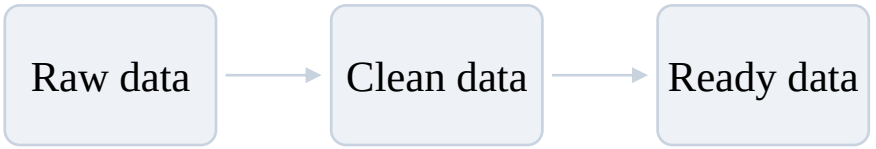
- Талдау нақты сұрақтан басталады: қандай шешім қабылдау керек?
- Дерек жинауда дерек көзі, өлшем бірлігі және уақыт аралығы белгіленеді.
- Тазалау сапасыз, қайталанған немесе жетіспейтін мәндерді өңдейді.
- Нәтиже түсіндіруге болатын қорытынды және әрекет ұсынысы болуы тиіс.

Дерек түрлері және дерек көздері

Дерек түрі	Сипаттамасы	Мысал	Талдау тәсілі
Сандық	өлшенетін мән	баға, уақыт	орташа, тренд
Категориялық	топ немесе класс	қала, өнім	жиілік, үлес
Уақыттық	уақыт бойынша реттелген	күндік сатылым	динамика
Мәтіндік	табиғи тіл	пікір, құжат	тақырып, эмоция
Кеңістіктік	орынмен байланысты	аймақ, координата	карта

Деректерді тазалау

• Жетіспейтін мәндер медиана, мода немесе арнайы белгі арқылы өңделеді.
• Қайталанған жазбалар бірегей кілт арқылы тексеріледі.
• Қате форматтар: күн, валюта, сандық өлшем және мәтін бір стандартқа келтіріледі.
• Аномалиялар бөлек қаралады: олар қате де, маңызды белгі де болуы мүмкін.



Тазалау сапасы талдау нәтижесіне тікелей әсер етеді: қате дерек дұрыс әдісті де қате қорытындыға әкеледі.

2. Статистикалық сипаттамалар және дерек сапасы

Бұл бөлімде деректің орталық мәні, таралуы, байланысы және сенімділігін бағалау үшін қолданылатын негізгі көрсеткіштер беріледі.

Сипаттамалық статистика

Көрсеткіш	Мағынасы	Қашан пайдалы
Орташа мән	барлық мәндердің ортасы	симметриялы деректе
Медиана	орталықтағы мән	ауытқу көп болса
Мода	ең жиі кездесетін мән	категориялық деректе
Дисперсия	мәндердің шашырауы	тұрақтылықты бағалау
Стандартты ауытқу	орташадан орташа қашықтық	тәуекел, өзгергіштік

Таралу мысалы

A

B

C

D

E

Таралу, тренд және маусымдылық

• Таралу мәндердің қай аралықта және қандай жиілікпен кездесетінін көрсетеді.
• Тренд көрсеткіштің ұзақ мерзімді өсу немесе төмендеу бағытын сипаттайды.
• Маусымдылық белгілі кезең сайын қайталанатын үлгіні білдіреді.
• Орташа мән жалғыз өзі деректі толық сипаттамайды.

Тренд үлгісі

Корреляция және себеп-салдар

Оң корреляция

• Корреляция екі айнымалының бірге өзгеру бағыты мен күшін сипаттайды.
• Корреляция себепті дәлелдемейді: үшінші фактор әсер етуі мүмкін.
• Себептік қорытынды үшін эксперимент немесе қосымша бақылау қажет.
• Визуализация бастапқы байланыстарды көруге көмектеседі.

Жетіспейтін мәндер және аномалиялар

Мәселе	Мүмкін себеп	Өңдеу жолы
Missing value	өлшеу болмады	толтыру немесе белгілеу
Duplicate	қайта тіркеу	біріктіру немесе жою
Outlier	шынайы ерекше жағдай немесе қате	IQR, z-score, тексеру
Inconsistent format	әртүрлі формат	бір стандартқа келтіру
Sampling bias	іріктеу тең емес	іріктеуді қайта қарау

3. Визуализация негіздері

Бұл бөлімде деректі көрнекі түрде түсіндіру қағидалары, график түрін таңдау және визуалды қателерден сақтану қарастырылады.



Жақсы визуализация қағидалары

• График бір нақты сұраққа жауап беруі керек.
• Ось атауы, өлшем бірлігі және уақыт аралығы түсінікті беріледі.
• Артық безендіру негізгі ойды әлсіретеді.
• Маңызды айырмашылық көзге көрінуі тиіс, бірақ мәндер бұрмаланбауы керек.
• Визуализация қорытындыны ауыстырмайды, оны дәлелдейді.



График түрін таңдау

Мақсат	График түрі	Мысал
Уақыт бойынша өзгеріс	line chart	айлық сатылым
Категория салыстыру	bar chart	өнімдер бойынша табыс
Үлес көрсету	pie / donut	нарық үлесі
Байланыс көру	scatter plot	баға мен сұраныс
Таралу көру	histogram / boxplot	бағалар шашырауы
Аймақтық дерек	map	өңірлік көрсеткіш

Дашборд құрылымы

KPI 1
Сатылым

KPI 2
Пайда

KPI 3
Өсім

Тренд

Салыстыру
A
B
C

Кесте / сүзгі / түсіндірме аймағы

Жоғарыда KPI, ортасында тренд пен салыстыру, төменде егжей-тегжей және сүзгі орналасса, есеп оқу ыңғайлы болады.

Визуализациядағы жиі қателер

Қате	Нәтижесі	Дұрыс тәсіл
Осьті нөлден бастамау	айырмашылық асыра көрінеді	масштабты көрсету
Тым көп категория	оқу қиындайды	TOP-N немесе топтау
Түсіндірмесіз график	негізгі ой жоғалады	атау және қорытынды
3D диаграмма	мән бұрмаланады	2D график
Түс көптігі	назар шашырайды	бірізді логика

4. Құралдар мен практикалық қолдану

Бұл бөлімде деректерді талдау мен визуализациялауда жиі қолданылатын құралдар, жұмыс ағыны және нәтижені ұсыну тәсілдері жинақталады.

Құралдар және олардың рөлі

Құрал	Күшті жағы	Қолдану орны
Excel	жылдам кесте, pivot	шағын есеп
Python	автоматтандыру, ML	зерттеу, модельдеу
SQL	дерекқор сұранысы	ETL, біріктіру
Power BI	интерактивті дашборд	бизнес есеп
Tableau	визуалды аналитика	презентациялық есеп

Практикалық жұмыс ағыны



- Дерек алдымен құрылымдалып, сапасы тексеріледі.
- Алғашқы талдауда бос мән, үлестіру және негізгі статистика қаралады.
- Нәтиже визуализация арқылы түсіндіріледі.
- Қорытындыда нақты шешім ұсынысы беріледі.

BI есеп бетінің дизайны

Бөлік	Мазмұн	Мақсаты
KPI қатары	сатылым, пайда, өсім	жылдам баға
Тренд графигі	уақыт бойынша өзгеріс	динамиканы көру
Салыстыру графигі	категория көрсеткіші	басым бағытты табу
Карта/сегмент	аймақ немесе топ	айырмашылықты көрсету
Сүзгі	жыл, аймақ, өнім	интерактивтілік

Аналитикалық қорытындыны ұсыну

- Бірінші сөйлемде басты нәтиже нақты айтылады.
- Сандар салыстырумен беріледі: жоспармен, өткен кезеңмен немесе орташа мәнмен.
- График дәлел ретінде қолданылады; түсіндірме қысқа және нақты болуы керек.
- Ұсыныс әрекетке бағытталады: не істеу керек, кімге, қашан және қандай нәтиже күтіледі.
- Шектеулер ашық жазылады: дерек толықтығы, мерзімі, болжам дәлдігі.



Қысқа кейс: сатылым дерегін талдау

Мақсат

Айлар бойынша сатылым динамикасын және тиімді өнім топтарын анықтау.

Дерек

Invoice, Date, Product, Region, Sales, Profit бағандары бар кесте.

Талдау

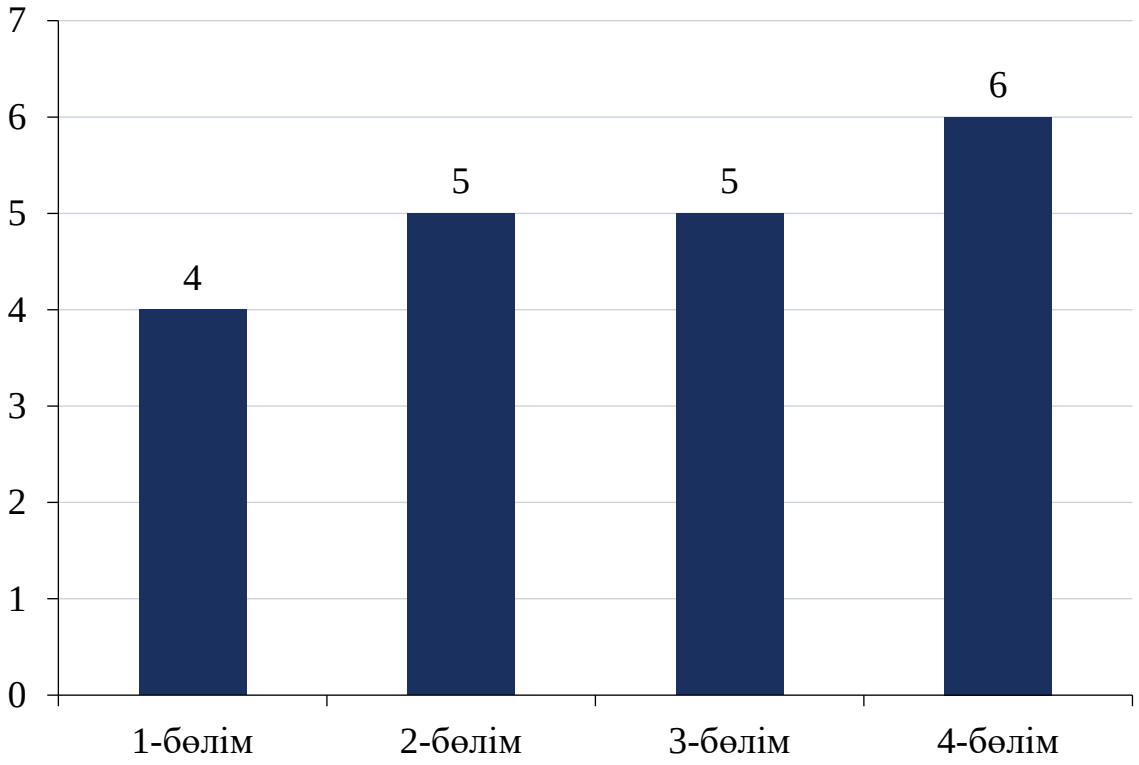
Тазалау, айлық тренд, өнім топтары бойынша салыстыру, пайда маржасы.

Нәтиже

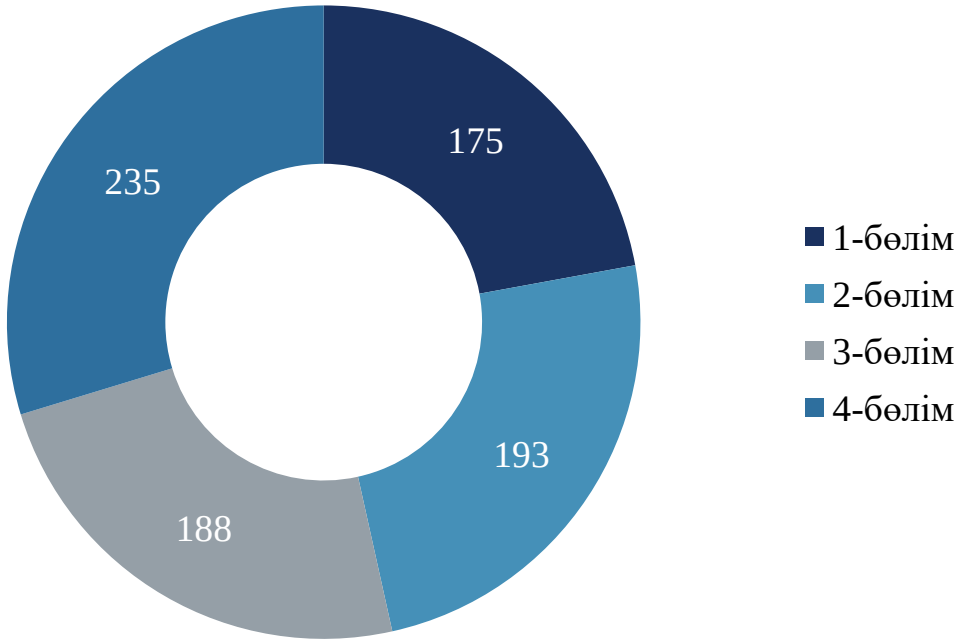
Маусымдық өсім байқалады; кей өнімдер сатылымы жоғары, бірақ маржасы төмен.

Диаграммалар

Бөлімдер бойынша слайд саны



Бөлімдер бойынша мәтін көлемі (сөз саны)



Қорытынды

1. Деректерді талдау нақты сұрақтан және сенімді дерек дайындаудан басталады.
2. Сипаттамалық статистика деректің орталық мәнін, таралуын және сапасын түсіндіреді.
3. Визуализация аналитикалық ойды дәл жеткізу құралы болып табылады.
4. График түрі дерек типіне және сұраққа байланысты таңдалады.
5. Жақсы есеп сан, график және қысқа түсіндірмені біріктіріп, шешім қабылдауға көмектеседі.

Негізгі ой: сапалы талдау - дұрыс сұрақ, таза дерек, түсінікті визуализация және нақты ұсыныс бірлігі.

Деректерді талдау және визуализациялау негіздері

Аналитикалық ойлау, сапалы дерек және түсінікті визуалды коммуникация
Дәріс 6