М.ӨТЕМІСОВ АТЫНДАҒЫ БАТЫС ҚАЗАҚСТАН УНИВЕРСИТЕТІ
ЖАРАТЫЛЫСТАНУ ГЕОГРАФИЯ ФАКУЛЬТЕТІ
БИОЛОГИЯ, ГЕОГРАФИЯ, ХИМИЯ БІЛІМ БЕРУ БАҒДАРЛАМАЛАРЫ

Білім беру бағдарламасы: 6В01509 – Биология
ПӘН: «Биометрия»

Дәріс 25

Тақырыбы: Кластерлік талдау және оның биологияда қолданылуы.

Лектор:
А.ш.ғ.к., доцент
Альжанова Багдагул Сактагановна

ДӘРІСТІҢ МАҚСАТЫ

Мақсаты

Кластерлік талдау әдісінің биологиядағы қолданылуын зерттеу, оның артықшылықтары мен мүмкіндіктерін көрсету.

Күтілетін нәтиже

Студент кластерлік талдаудың мақсатын және негізгі ұғымдарын және дендрограмманы оқып, кластерлердің қалыптасуын түсіндіреді, кластерлік талдауды биологиялық және экологиялық деректерге қолдану жолдарын ұсынады.

.

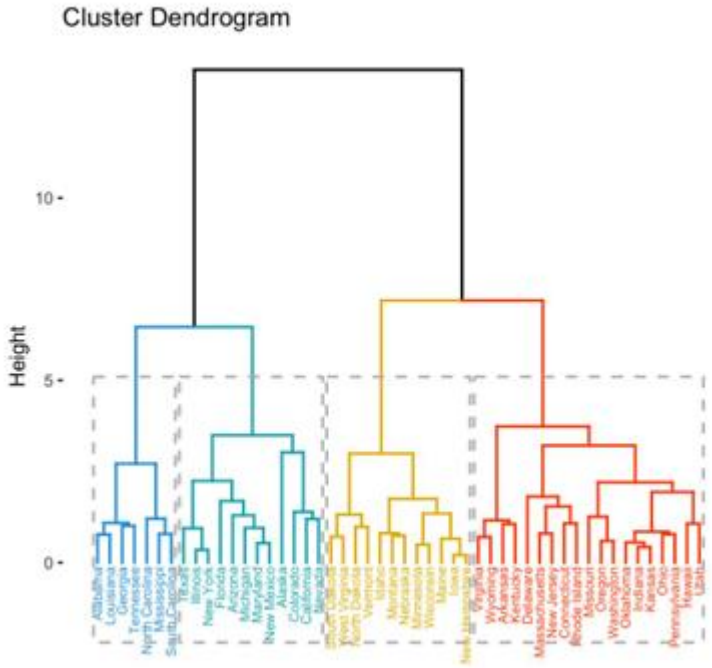
ДӘРІСТІҢ ЖОСПАРЫ

1. Кластерлік талдаудың мәні, мақсаты және негізгі ұғымдары.
2. Кластерлік талдаудың негізгі әдістері және оны жүргізу кезеңдері
3. Негізгі компоненттер әдісі: көпөлшемді деректерді ықшамдау және визуализациялау.
4. Кластерлік талдау мен негізгі компоненттер әдісінің биология және экологияда қолданылуы.

Кластерлік талдаудың мәні, мақсаты және негізгі ұғымдары

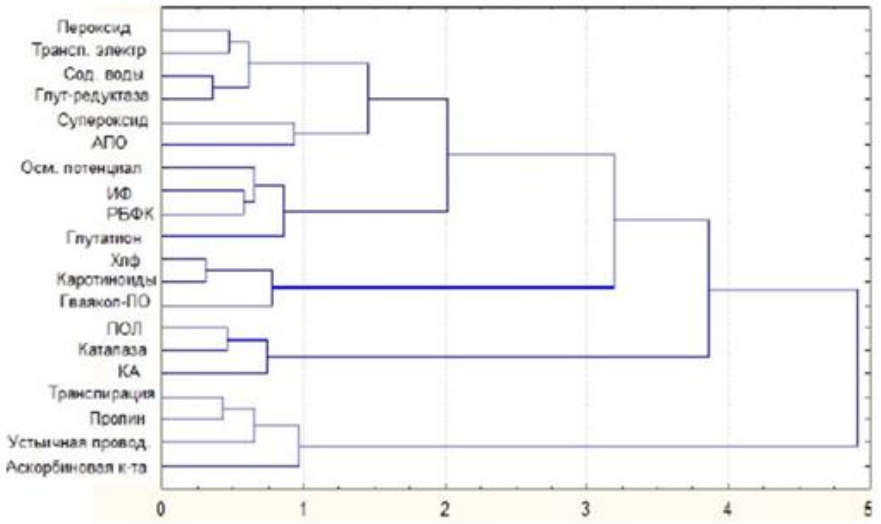
Кластерлік талдау – зерттелетін нысандарды олардың белгілері бойынша салыстырып, өзара ұқсас нысандарды бір топқа, яғни кластерге біріктіретін көпөлшемді статистикалық әдістер жиынтығы.

Кластер ішіндегі нысандар бір-біріне барынша ұқсас, ал әртүрлі кластерлерге кіретін нысандар бір-бірінен барынша ерекшеленуі қажет.



Кластерлік талдаудың мәні, мақсаты және негізгі ұғымдары

Кластерлік талдаудың негізгі мақсаты – алдын ала белгіленген топтар болмаған жағдайда деректердің табиғи құрылымын анықтау. Сондықтан ол бақылаусыз оқыту әдістерінің қатарына жатады. Кластерлеуде зерттеу нысандары даралар, популяциялар, түрлер, өсімдік қауымдастықтары, су айдындары, гендер немесе биологиялық сынамалар болуы мүмкін.



Кластерлік талдау объектілерді олардың айырмашылығына немесе ұқсастығына қарай санатқа топтастыру үшін қолданылатын бірнеше көп өлшемді (көп вариантты) әдістерді қамтиды. Алгоритм-бір жағынан топ ішіндегі вариацияларды азайту және оларды екінші жағынан топтар арасында көбейту.

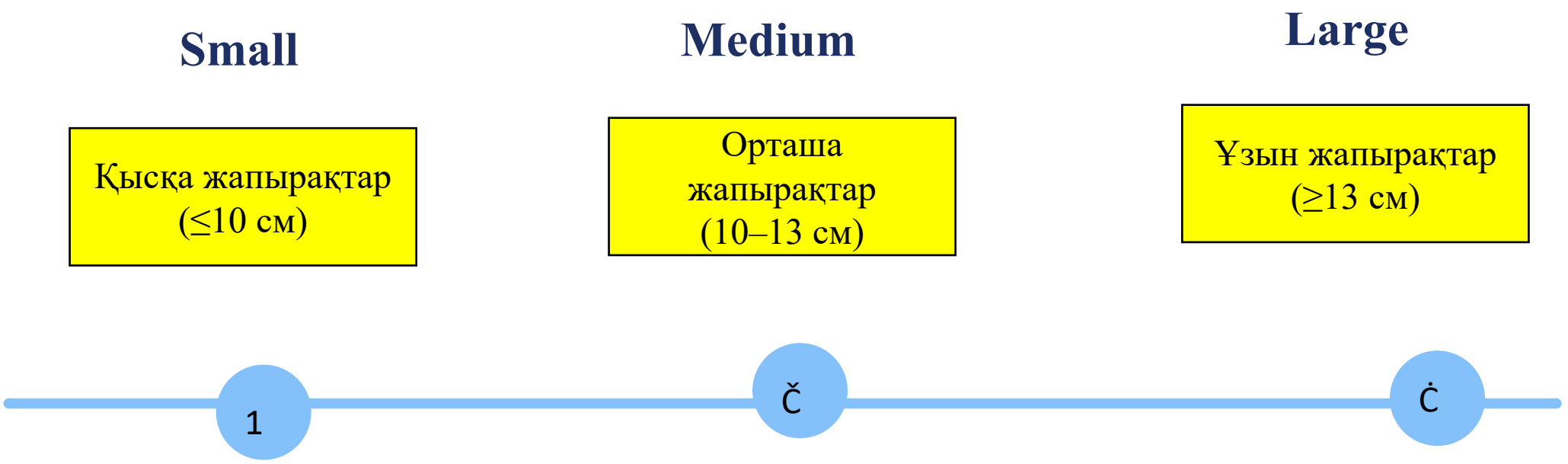
Кластерлік талдаудың негізгі артықшылығы

Алдын ала қатаң гипотеза талап етпей, деректердің ішкі құрылымын өздігінен анықтай алу қабілетінде.

Бұл қасиет оны әсіресе зерттеу объектілері арасындағы байланыстар толық анықталмаған жағдайларда тиімді құралға айналдырады. Сонымен қатар, кластерлік талдау нәтижелері көрнекі түрде ұсынылатындықтан (мысалы, дендрограмма, кластерлік карталар), зерттеу нәтижелерін түсіну мен талдау процесі жеңілдейді.

Кластерлік талдаудың негізгі әдістері және оны жүргізу кезеңдері

Жапырақтар ұзындығына қарай 3 кластерге бөлінді, бұл табиғи популяция үшін қалыпты жағдайды көрсетеді.



Негізгі компоненттер әдісі: көпөлшемді деректерді ықшамдау және визуализациялау.

Негізгі компоненттер әдісі, яғни Principal Component Analysis – PCA, өзара байланысқан көптеген сандық айнымалыларды ақпараттың басым бөлігін сақтай отырып, аз санды жаңа компоненттерге түрлендіреді. Әрбір негізгі компонент бастапқы айнымалылардың сызықтық комбинациясы болып табылады.

Бірінші негізгі компонент деректердегі өзгергіштіктің ең үлкен бөлігін түсіндіреді. Екінші компонент бірінші компонентке тәуелсіз болады және қалған өзгергіштіктің мүмкіндігінше көп бөлігін қамтиды. Нәтижесінде ондаған белгіні екі немесе үш негізгі компонент арқылы көрнекі түрде көрсетуге болады.

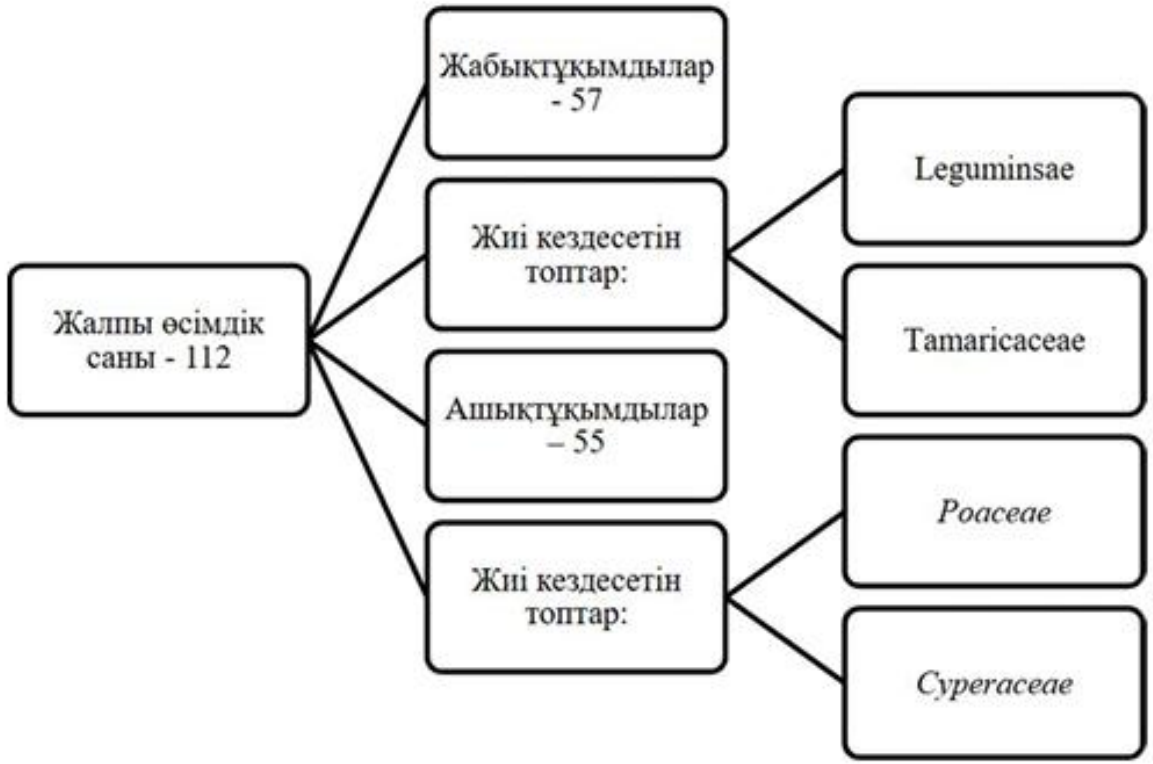
Кластерлік талдау мен негізгі компоненттер әдісінің биология және экологияда қолданылуы.

Кластерлік талдау биологияның әртүрлі салаларында күрделі деректерді топтастыру үшін қолданылады. Таксономияда түрлер мен популяциялар морфологиялық, анатомиялық немесе генетикалық белгілері бойынша салыстырылады. Генетикада гендердің экспрессиялық белсенділігі ұқсас гендер бір кластерге біріктіріледі. Микробиологияда микроағзалар биохимиялық және молекулалық белгілеріне қарай топтастырылады.

Мысалы, 20 зерттеу учаскесіндегі өсімдік түрлерінің кездесуі мен проективтік жамылғысы тіркелсе, кластерлік талдау флоралық құрамы ұқсас учаскелерді біріктіреді. Нәтижесінде шалғындық, далалық, шөлейттік немесе антропогендік өзгеріске ұшыраған қауымдастықтар ажыратылуы мүмкін.

Мысал 1

Зерттеу жұмысы негізінде жалпы өсімдіктердің 112 түрі тіркелді. Олардың 57-сі жабықұқымдылар, 55-і ашықұқымдылар



Жоғарыда көрсетілген мәліметтер бойынша 112 өсімдік түрі зерттеліп, олардың кластерлік топтары анықталды. Кластерлік талдау өсімдіктерді олардың тұқымдық ерекшеліктері мен жиілік көрсеткіштеріне қарай топтастыруға мүмкіндік береді

Экология ғылымында кластерлік талдау әртүрлі экожүйелердегі организмдерді, әсіресе өсімдіктер мен жануарларды, олардың ұқсастық белгілеріне қарай топтастыру үшін кеңінен қолданылады

Экологиядағы кластерлік талдаудың негізгі бағыттары:

1 Қауымдастықтарды жіктеу: өсімдіктер немесе жануарлар қауымдастықтарын ұқсастықтарына қарай топтау.

2 Биоәртүрлілікті бағалау: әртүрлі экожүйелердегі түрлердің таралуын және олардың әртүрлілігін салыстыру.

3 Қоршаған орта факторларының әсерін зерттеу: экожүйелердегі түрлердің экологиялық бейімделуін бағалау

ҚОРЫТЫНДЫ

Кластерлік талдау қазіргі таңда көпөлшемді деректерді өңдеу мен талдаудың маңызды әдістерінің бірі ретінде ғылыми зерттеулерде кеңінен қолданылып келеді. Әсіресе биология ғылымында деректердің күрделілігі мен көлемінің артуына байланысты оларды жүйелеу, құрылымдау және интерпретациялау қажеттілігі оғарылаған сайын кластерлік талдаудың маңызы арта түсуде. Бұл әдіс зерттелетін объектілерді олардың ұқсастық белгілеріне қарай топтастыру арқылы деректер арасындағы жасырын заңдылықтарды анықтауға мүмкіндік береді..

ПЫСЫҚТАУ СҰРАҚТАРЫ

Кластерлік талдаудың негізгі мақсаты қандай?
Кластер мен айнымалы ұғымдарының айырмашылығы неде?
Иерархиялық және иерархиялық емес кластерлеу қалай ерекшеленеді?
Дендрограмма қандай ақпарат береді?
Деректерді кластерлеу алдында не үшін стандарттау қажет?

ҚОЛДАНЫЛАТЫН ӘДЕБИЕТТЕР ТІЗІМІ

1) Аубакиров Х.Ә. Биометрия: оқулық / Х.Ә.Аубакиров. – А.: Дәуір, 2011. – 165-176 б.; 337-339 беттер.

2) Раманқұлова А.А. Биологиялық статистика: оқу құралы. – Алматы: Эверо, 2023. – 169-187 бет.

3) Шоқаманов Ю.К., Белгібаева Қ.Қ. Статистика . – Алматы: Экономика. – 540 бет.

4) Хили Д.Ф. Статистика негіздері: әлеуметтік зерттеу құралы. – Алматы : «Ұлттық аударма бюросы» қоғамдық қоры, 2020. – 576 бет.

5) Биологиядағы бағдарламамен қамтамасыз етілген статистикалық әдістер: оқу-әдістемелік құралы / А.М. Калимағамбетов, З.М.Бияшева, А.В. Ловинская, С.Б. Дәулетбаева. – Алматы: Қазақ университеті, 2019. – 58-64 б.

Назарларыңызға рахмет!

