



ADATÉRTELMEZÉS- ÉS KOMMUNIKÁCIÓ

A DUCK tanfolyam ismeretanyagának összefoglalója

TARTALOMJEGYZÉK

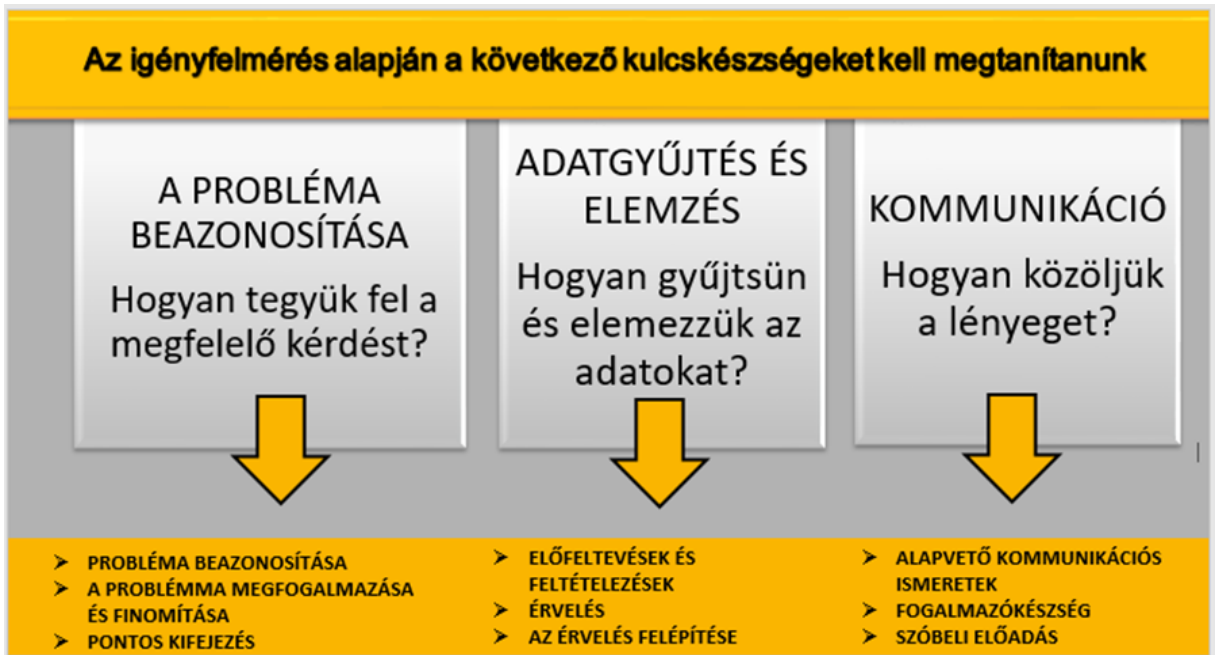
BEVEZETŐ	2
1. MODUL: PROBLÉMAELEMZÉS	3
1. KÉSZSÉG: A PROBLÉMA MEGFOGALMAZÁSA	3
2. KÉSZSÉG: KÉRDÉSEK MEGFOGALMAZÁSA ÉS FINOMÍTÁSA	4
3. KÉSZSÉG: KIFEJEZÉSEK ÉS FOGALMAK PONTOS DEFINIÁLÁSA	9
2. modul: adatelemzés	11
1. KÉSZSÉG: ADATÉRTELMEZÉS	11
2. KÉSZSÉG: A STATISZTIKÁK ÉS A VALÓSZÍNŰSÉG MEGÉRTÉSE	12
3. KÉSZSÉG: HIPOTÉZISEK ÉS ELŐFELTEVÉSEK	14
4. KÉSZSÉG: ÉRVELÉS	14
5. KÉSZSÉG: AZ ÉRVELÉS FELÉPÍTÉSE	15
3. MODUL: DATA COMMUNICATION	18
1. KÉSZSÉG: ALAPVETŐ ADATKÖZMUNIKÁCIÓS KÉSZSÉGEK	18
2. KÉSZSÉG: FOGALMAZÓ KÉSZSÉG	19
3. KÉSZSÉG: SZÓBELI BEMUTATÓI KÉSZSÉGEK	21

A különböző iparágak egyre több adatot generálnak, ezzel foglalkoznak többek között a dolgok internetje illetve a big data elméletek, és ez az adathalom erős analitikus készségeket kíván meg a legkülönbözőbb területeken dolgozóktól. Az adatok értelmezése és ennek a kommunikációja kulcsfontosságúvá vált a cégek, intézmények számára, és az adattudós (data scientist) egyre keresettebbé válik. Az adatok értelmezésének és vizuális ábrázolásának készsége (pl. infografika készítése), az adatok elemzésének és bemutatásának készsége keresetté vált a modern munkaerőpiacon, ezeknek a hiánya komoly problémát jelent valamennyi fél részére.

Ezen túlmenően, a tanárok és oktatók adatelemzési és adatkommunikációs készsége nem csak a kutatásaik és azok bemutatásának minőségi kérdése: nem csak szűkebb szakmai adatokat kell értelmezniük, de a hallgatóikra és azok igényeire vonatkozó adatokat is.

A DUCK projekt egy 30 órás multimédiás, kevert módszerű tananyagot dolgozott ki, mely három modulból: problémaelemzés, adat elemzés, adatkommunikáció modulokból. A tananyag a szakképzésben, a felsőoktatásban és a felnőttképzésben való bevezetésre is alkalmas. Elvégzését követően a tanárok és oktatók (valamint tanulók) képesekké válnak hatékonyabban elemezni és prezentálni az adatokat, ezáltal gazdagodik az oktatásuk minősége és javulnak az elhelyezkedési lehetőségeik is.

Jelen dokumentum összefoglalja a tananyag legfontosabb tartalmi részeit. Alkalmas jegyzetként való használatra a tanulás során, illetve emlékeztetőként a kurzus elvégzését követően.



1. MODUL: PROBLÉMAELEMZÉS

Tanulási célok:

Mi a kérdés pontosan?

- Vizsgálja meg a problémát a lehetséges célok szempontjából.
- A kérdés megfogalmazása és finomítása: ismerje meg a különféle típusú kérdéseket, a jó kérdés főbb jellemzőit, fordítsa át a kérdést problémává.
- Mi áll a rendelkezésünkre? - az inputok megértése: adatok, fogalmak, attribútumok, példák.
- A rendelkezésre álló információ lehetséges megjelenítési formái a megoldás szempontjából: táblázatok, lineáris modellek, fadiagramok, klaszterek.

1. KÉSZSÉG: A PROBLÉMA MEGFOGALMAZÁSA

Az alapvető beazonosítása a valódi problémának azzal kezdődik, hogy elemezzük azt a levelet, üzenetet, vagy szöveget, amelyhez a következőképpen juthattunk:

1. A vezető konkrétan megadja a címben: Termelési válság – hogyan reagálunk
2. A probléma a szöveg elején tisztán, jól kivehetően le van írva.
3. A szövegben található valami fogódzó: „Hogy reagáljon a rektor ilyen szituációban?”

4. A háttér ismerete, vagy a szöveg megalkotójának felelősségköre segít bennünket: A tudományos rektorhelyettes inkább a kutatási adatok növekedését fogja inkább keresni inkább, mint a beiratkozási vagy bevételi adatokat.

Sok esetben a helyzet kevésbé egyszerű, a probléma rejtett. A probléma megtalálásához néha az eligazító szöveg végkövetkeztetését kell elemeznünk, mely előre jelzi a feltételezett megoldását a problémának. Tehát ha megtaláljuk a konklúziót, ebből a valós kérdésre vagy problémára következtethetünk. A végkövetkeztetés megtalálásához:

1. Indikátor szavakat kereshetünk, mint például: tehát, így, következésképp.
2. Tipikus helyek megnézése: Üzenet, szöveg eleje, vége.
3. Kizárás: kizárjuk a példákat, statisztikákat, definíciókat, bizonyítékokat, melyek nem konklúziók.
4. Utánanézzük a szerző személyének.

Vannak leíró és előíró kérdések.

A leíró kérdések estén a tipikus kérdések a „Hol?”, „Mi?”, „Mikor?” és „Hogyan?” kérdőszavakkal kezdődnek.

A tipikus előíró kérdések a következőkkel kezdődnek: „Kellene?”, „Kell?”, „Muszáj”? Jó ez így?”.

Először meg akarjuk érteni, miért történnek a dolgok (szerkesztés), és csak egy második lépésben azt, vajon ez jó vagy rossz a megfigyelő számára.

2. KÉSZÉG: KÉRDÉSEK MEGFOGALMAZÁSA ÉS FINOMÍTÁSA

Kérdéstípusok:

1. **Egyszerű leíró kérdés.** Ez a kérdés típus a legalapvetőbb a hatból. Ez a leginkább tényszerű, és legkönnyebben megérthető. Ez a kérdés összegezi, kiemeli egy adott adatsor, adathalmaz jellemzőit. Nincs benne értelmezés, az egyszerű leíró kérdés tényekhez vezet el bennünket. Néhány példa: Az egyetemre beíratkozottak számának átlaga utolsó tíz évében; a nők aránya egy adott évben az iskolában; a munkaórák átlaga egy vállalat adott periódusában.
2. **Feltáró (felderítő) kérdés.** Ez egy hipotézis létrehozó kérdés egy általános megérzés vagy ötlet alapján, hogy ellenőrizzünk, trendeket, mintázatokat, összefüggéseket. Tehát például, ha úgy érezzük, hogy a vizsgaeredmények valamimódon összefüggésben lehetnek az alvási idővel, vagy ha a lemorzsolódási arányról úgy érezzük, hogy valamimódon a vizsga időzítési

mintázatokkal lehet összefüggésben, akkor feltáró kérdést teszünk fel. Az első példánk esetét véve egy feltáró elemzést végezhetünk az alvási idő és a vizsga eredmények kapcsolatáról. Ha úgy találjuk, hogy bizonyos alvási időintervallumok jobb vizsgajegyeket mutatnak, mint azoknál, akik kevesebbet alszanak, egy hipotézist állíthatunk fel, mondjuk hogy: „A hallgató nyolc óra alvása a vizsganap előtt kevesebb bukással járó vizsgához vezet.”

3. **Következtetési (pontosító) kérdés.** Ez a kérdés az eredetinek egy újra fogalmazása, vagyis ismételt kimondása a hipotézisnek, hogy megtudjuk, pontosan mely adathalmazokra és milyen mértékben igaz a hipotézisünk. A mi példánkban elemezhetjük, hogy az alvási idő különböző egyetemek különféle tantágyára is igazak-e, és megvizsgálható, hogy igaz-e a hipotézis a saját országunkban vagy Európában az összes egyetem valamilyen képzésére. Szerencsés esetben reprezentatív adatunk volt Európa minden felsőoktatási intézményéről, így az eredeti hipotézisünkből következtethetünk Európa minden egyetemére.
4. **Prediktív (jósló) kérdés.** A prediktív kérdés azt akarja feltárni, hogy egy adott viselkedés vagy tény milyen körülmények között, figyelhető meg. Azaz a prediktív kérdés azt próbálja meg feltárni, hogy mi az, ami megjósol bizonyos tényeket vagy viselkedéseket. A korábban említett példánkat felhasználva a nyolc óra alvást megjósolhatja (előre jelezheti) önálló lakáskörülmények, vagy például fő tevékenységként folytatott tanulás a munka mellett való tanulással szemben.
5. **Ok-okozati (kauzális) kérdés.** Az ok-okozati kérdés az egyes faktorok közti kapcsolatra kérdez rá. Arra kérdez rá, hogy emelve vagy csökkentve egy tényezőt, hogyan változtat meg egy másik tényezőt. A mi példánkban megkérdezhetjük, hogy mi történik a vizsgajegyekkel, ha emeljük vagy csökkentjük az alvási időt. Ez a kérdés például arra próbál választ találni, hogy a több alvás jobb vizsgaeredményeket eredményez-e. (Még mindig a példánknál maradva könnyen lehet, hogy egy bizonyos mennyiségű alvás után már újra visszaesnek az eredmények.)
6. **Mechanikus kérdés.** Ez a kérdés típus azt próbálja meg felderíteni, hogy az egyik tényező miként, hogyan befolyásol egy másik tényezőt. A mi példánkban például feltehetjük a következő mechanisztikus kérdést: „Hogyan befolyásolja a vizsga előtti alvás a vizsgajegyeket?”

A három lépéses modell egy elég egyszerű megközelítése ennek:

1. Definiáljuk az elvárásainkat a kérdésre vonatkozóan;

2. Gyűjtsünk információt a kérdésről;
3. Döntsük el, hogy a gyűjtött információ kielégíti-e az elvárásainkat, és ha nem, pontosítsuk a kérdést, vagy elvárást.

Ez egy hurokká is válhat, ha a végén még mindig nem vagyunk megelégedve az eredménnyel. Kezdjünk egy új kört, új elvárásokkal vagy kérdésfeltevessel, és nézzük meg, hogy tudunk-e gyűjteni új információt a kérdésről, és hogy az információk megfelelnek-e az újabb elvárásoknak

	Kérdés 1. változat	
Elvárások	Összehasonlítás: megfelel az elvárásoknak?	Információ a kérdés 1. változatáról
Új elvárások	Kérdés 2. változat	Információ a kérdés 2. változatáról

Milyen információk lehetnek fontosak a kérdésfeltevessel kapcsolatban? A következőkben felsorolunk öt olyan szempontot, amire fontos lehet gondolni:

1. A kérdésnek a döntéshozók és más érdekelték részére érdeklődésre számot tartónak, azaz érdekesnek kell lenni az adott problématerületen. Ők az elsődleges és másodlagos célcsoportunk. Ha a beiratkozási szabályok 35 évben limitálják a felvehető korát, akkor az ötvenéves hallgatók vizsgaeredményeinek adatai nem érdekesek ebben a kontextusban.
2. A kérdést még nem válaszolták meg. Egyre több publikus adat érhető el a világhálón. Nagyon is lehetséges, hogy a szakirodalomban a kérdésünkre már megvan a válasz. De a szakértőktől való információ gyűjtés során (2. lépés) az is lehetséges, hogy egy hasonló, vagy a mienkkel kapcsolatos kérdésre van válasz, eredmény. Ez átalakíthatja a kérdésfelvetésről, vagy elvárásokról alkotott eredeti elképzelésünket. Amennyiben például egy új on-line anyagot szeretnétek készíteni a beiratkozás előtt a felvételizőknek, nagyon valószínű, hogy a kar már készített felmérést a célcsoporttal az on-line tanulási érettségtől eltérő okokból. Ennek ellenére, ha sikerül megtudni az okostelefonnal rendelkezők arányát, és a Youtube használati adataikat, lehet, hogy ezek segítenek megválaszolni azt a kérdést, hogy mennyire várható, hogy ugyan ezeken a médiumokon keresztül tanuljanak is.
3. A kérdés életszerű, megválaszolható. Ez egy gyors ellenőrzés, hogy a kérdés logikailag két különböző témához tárgyhoz, vagy egymáshoz kapcsolható témához, tárgyhoz tartozik saját tapasztalatunk szerint. A mérnöki tanulmányokra való jelentkezés összefüggése a szülők mérnöki hátteréhez, egy természetesen felmerülő, elképzelhető összefüggés, míg a tanulók

színpreferenciájának összefüggése nem tűnik életszerűnek, kivéve, ha erre valamiféle előzetes tudás utal.

4. A kérdés megválaszolhatatlan. Vannak nagyon jó kérdésvetések, melyek valójában megválaszolhatatlanok a mérhetetlenségük, forráshiány vagy erkölcsi akadályok miatt. Az informatikai érdeklődés összefüggése a játékfüggőséggel megválaszolhatatlan kérdés, mivel a függőség érzékeny (minősített személyes) adat, melyen legálisan nem lehet gyűjteni engedély nélkül.
5. A kérdés nem elég specifikus (azaz túl általános vagy homályos). Az olyan kérdések, mint szebb, erősebb, boldogabb, elégedettebb, mind pontosítandók. Hogyan mérjük a boldogságot? A jobb, gyorsabb, több és más összehasonlítások sem elég specifikusak. Meg kell pontosan határoznunk, hogy milyen mértékben akarunk látni különbséget két mérés között. (kicsit gyorsabb, vagy 10%-al gyorsabb, vagy 5 km/h-val gyorsabb)

További jellemzők:

1. A kérdésnek operacionalizáltnak kell lennie. Ez azt jelenti, hogy a kérdést lefordíthatjuk egy annak megfelelő adatgyűjtéssé. Ha a kérdésünk nem eléggé operacionalizált, azaz konkretizált, akkor az összegyűjtött adathalmaz többféle értelmezésre adhat majd okot. Ennek elkerülésére a begyűjtött adatoknak ésszerűen specifikusnak kell lennie azokból a szempontokból melyek segítségével a kérdés megválaszolható. Például: Ha azt akarjuk megtudni, hogy a hallgatók hőérzetét kánikula esetén vízpára permetezéssel megoldhatjuk-e, és hogy ennek az indikátora a bekapcsolt ventilátorok száma lenne, megtévesztő lehet. Azért, mert a ventilátorok be- vagy kikapcsolását az osztályteremben további tényezők is befolyásolhatják, mint a huzat, vagy az okozott zaj.
2. A kérdésnek nem szabad zavaró elemet tartalmaznia. Ez egy lehetséges probléma amikor a kérdés két tényező közötti összefüggésre, kapcsolatra próbál választ találni. Előző példánkban a ventilátor zaja, vagy huzatja zavaró tényezők. Mindkettő befolyásolja a bekapcsolást és a vízpárásítást.
3. A kérdésvetés nem vezethet elfogult eredményhez. Tipikus elfogultságnak ismerjük a visszaemlékezési elfogultságot. Előző példánkat folytatva: amennyiben az egyetemen hirdetést adunk fel kánikula kutatásunkhoz, nagyon valószínű, hogy akik érzékenyek a melegre, hajlamosabbak jelentkezni a hirdetésre. Következésképp a memóriájuk elfogultabb lesz sok apró hőmérsékleti tényezővel kapcsolatban, például a ventilátorok zajával kapcsolatban.

Nyitott vagy zárt kérdés?

A kérdés zárt, ha csak igennel és nemmel válaszolhatunk. (Érti az Ohm törvényt? – Igen.) A kérdés szintén zártnak tekinthető, ha a kérdésfeltevésben a válaszadót megoldás orientáltan tesszük fel. Azaz az elvárt választ előjelezzük a kérdésben. (Meggkérnéd az osztálytársad, hogy megoldja ezt a feladatot? – Igen.) Egy nyílt kérdést természetesen nem lehet igennel vagy nemmel megválaszolni.

A nyílt és zárt kérdéseknél maradva érdemes tudni, hogy kutatási helyzetben amennyiben tényeket és problémákat akarunk feltárni mindig nyílt kérdéseket kell feltennünk. Amennyiben egy kérdést pontosítani kell egy választ, vagy le akarunk kerekíteni egy interjút, illetve be akarjuk fejezni, zárt kérdéseket érdemes feltennünk.

Egy másik vonatkozása mások „kikérdezésének” adatgyűjtés során az a tény, hogy versengő (erőforrás elosztási probléma) vagy együttműködő (kollégákkal történő beszélgetés) szituációban találhatjuk magunkat. Az adott helyzetek tipikus kihívásai a következők:

- **Versengő** kutatási megbeszélések kihívása: A kolléga vonakodik az információt megosztani, akár hazudik is. Lehetséges megoldás: Használjunk zárt kérdéseket hogy elkerüljük a kitérő válaszokat. Használjunk részletező rákövetkező kérdéseket (még akkor is, ha redundánsak) hogy kicsikarjunk némileg több információt.
- **Együttműködő** kutatási megbeszélések kihívása: A barátságos kollégák távol tartják a konfliktust vagy vonakodnak rossz hírt közölni. Lehetséges megoldás: Használjunk nyílt kérdéseket (Ha az ördög ügyvédje lennél, mit mondanál?) hogy kihúzzuk belőlük a negatív visszajelzést is. Kezdjük a legkevésbé érzékeny kérdéssel, és kapcsolatot építve lépcsőzetesen haladjunk a kínosabb kérdések felé.

Ebben a modulban, mely a probléma megfogalmazásáról (adatértelmezésről) szól, fontos megfogalmazni, hogy:

1. Nagyon specifikus problémák esetén zárt kérdésfeltevést javasolunk. (Mondhatjuk hogy... Jól értjük, hogy..., Egyetért azzal, hogy...)
2. Feltáró probléma esetén nyílt kérdéseket javasolunk. (Hogyan, mikor? Hol? Ki?)

3. KÉSZSÉG: KIFEJEZÉSEK ÉS FOGALMAK PONTOS DEFINIÁLÁSA

Kevesebb szó használata különösen hasznos munkakörnyezetben, amikor a főnökség tiszta és tömör jelentéseket és ajánlásokat vár el. A tiszta írás (és beszéd) elvárásának úgy tudunk eleget tenni, ha megtanuljuk a következőket:

- **Pontos nyelvezet használata:** Pontos, precíz főnevek és élénk igék használata segít erős képiséget teremteni és a terjengősséget elkerülni.
- **Tömör nyelvezet használata:** A lehető legkevesebb szó használata anélkül, hogy a jelentést feláldoznánk, a szöveget érthetőbbé teszi.
- A precíz nyelvezet tekintetében gyakran találjuk magunkat abban a helyzetben, hogy félrevezet bennünket az újdonság, újszerűség vonzása és hogy úgy tűnjön, hogy szakértők és profi szakemberek közösségekhez tartozóknak, akadémiai körökhöz tartozóknak tűnjünk. Fogadja meg, hogy sosem áldozza fel az érthetőséget az újszerűség és az identitás oltárán. Mielőtt olyan szinonimát használ, melyet nem pontosan ismer, nézzen utána a lehetséges mellékértelmezésnek (konnotációnak) is.
- Kerülje a szakzsargont. Még akkor is, ha a hallgatóság járatos benne, jobban érthető, ha egyszerű nyelvezetet használ.
- Definiálja a fogalmakat! Ha technikai jellegű fogalmakat használ, definiálja őket pontosan rögtön az elején. Amikor nehéz fogalmakkal élünk nagy a kísértés, hogy absztrakt, homályos, általános meghatározásokat alkalmazzunk. Ezek félreérthetőek, azaz homályosak. Amellett, hogy homályosak gyakran a parttalanul hosszú, ködös mondatok tengerében veszünk el. Az ellenkezőjét javasoljuk: legyen tömör! Konkrétan és tömörnek nem könnyű lenni.

A kétértelműség a nyelv egy természetes velejárója. Amikor valami vagy valaki kétértelmű, az nem feltétlenül jelent negatív dolgot, vagy helytelen cselekedetet.

Ennek ellenére amikor érvelünk azért, hogy mások elfogadják a következtetéseinket felelősek vagyunk azért, hogy tisztázzunk minden lehetséges kétértelműséget. Ez fordítva is igaz: amikor mások megpróbálnak befolyásolni bennünket, (reklámokban, politikában vagy amikor munkahelyen két lehetőség közül kell választanunk), az ő felelősségük, hogy tiszták legyenek azok a fogalmak melyeket használnak. Egész addig, amíg az érvelés kétértelmű, nem érdemes foglalkoznunk annak hasznosságával, értékelésével.

- Mivel az írott és beszélt nyelvünk tele van kétértelműséggel, ezért érdemes leszűkíteni a keresést: a leglényegesebb kifejezések és mondatok az érvelés végkövetkeztetésében és a végkövetkeztetés megindoklásában szoktak lenni.
- Amennyiben ezeket beazonosítottuk, nézzük meg, hogy a szerző adott-e definíciót a használt kifejezésekre. Amennyiben nem, a szövegkörnyezetet használhatjuk a jelentésük meghatározására. Például a „digitális” kifejezésnek több jelentése is lehet attól függően,

hogy oktatási környezetben, vagy egy biológiai konferencián bukkan fel: az egyik esetben egy elektronikus formátumról, másik esetben pedig az ujjak mozgatásának készségéről van szó.

- Amennyiben a szövegkörnyezet nem elégséges, és nincs módunk a szerzőnél rögvést tisztázni a jelentést, előre vihet az a módszer, ha megnézzük, mi történne, a kifejezések különböző definícióinak behelyettesítésével. Megváltoztatná-e a kimenetelt, azaz, hogy egyetértünk, vagy nem értünk egyet a szerzővel. Ha igen, az okfejtés hasznát nem tudjuk eldönteni.

Egy másik jó módszer lehet a kétértelműség behatárolásának, ha megkeressük az absztrakt szavakat és mondatokat az érvelésben vagy a végkövetkeztetésben. Az olyan szavak, mint egyenlőség, igény vagy hanyatlás több magyarázatot is takarhatnak, ezért valószínű, hogy további információra lesz szükségünk a tisztázáshoz.

Tanulási célok:

- Képesség az adatok gyűjtésére, tisztítására és elemzésére (különböző fájl típusok keresése; adatok kinyerése; az adatok rendezése, egyesítése és kezelése);
- statisztikák és valószínűség feltárása / megértése (adattípusok; valószínűség és várható értékek; változékonyság, eloszlás, aszimptotikus);
- megérteni az adatokat vizualizációval;
- megérteni az adatokat leíró statisztikákkal (bepillantani az adatokba, meghatározni azok dimenzióit és adattípusait, kiszámolni, megjeleníteni és megvitatni az egyszerű statisztikákat, és összefoglalni az adatokat);
- alapvető eszközöket használata az adatok elemzéséhez.

1. KÉSZSÉG: ADATÉRTELMEZÉS

Különbséget teszünk adatok, fogalmak, attribútumok között:

Az **adatok** megfigyeléseket tartalmaznak az adatbázisban található rekordokról, naplójegyzetokről, képekről a merevlemezen. Ezek a megfigyelések hasznosnak vagy haszontalannak, pontosnak vagy pontatlannak, hasznosnak vagy haszontalannak bizonyulhatnak. Kezdetben ezek csak megfigyelések. A megfigyelés önmagában nem elegendő a cselekvéshez. Amikor a megfigyeléseket összekapcsoljuk azzal, hogy a kérdésünk/kíváncsiságunk tárgya hogyan működik, akkor nyílik lehetőségünk tudás megszerzésére.

Fogalomról beszélünk, amikor formálisan definiáljuk a kutatás tárgyát. Ez egy adott példány általános, alapvető formája (konkrét, bár különböző).

Az **attribútum** egy objektum jellemzője vagy minősége (személy, dolog stb.), szorosan kapcsolódik a változókhoz. A változó az attribútumok logikai halmaza. A változók "változhatnak" - például lehetnek magasak vagy alacsonyak.

Grafikon készítése az adatok vizualizálása érdekében jó módszer arra, hogy jobban megértse a kérdését és adatait. A grafikai lehetőségek a bemutatni kívánt adatoktól függenek. Az adatmegjelenítés grafikontípusai a következők:

- Megoszlás: hegedű, sűrűség, hisztogram, boxplot, ridgeline;
- Korreláció: szórás, hő térkép, korrelogram, buborék, összekapcsolt szórás, sűrűség 2d;
- Rangsor: sáv, pók / radar, világfelhő, párhuzamos, nyalóka, kör alakú sáv;
- Az egész része: treemap, venn diagram, fánk, kördiagram, dendrogram, kör alakú csomagolás, napsugár;
- Evolúció: vonalvezetés, terület, halmozott terület, folyamatábra;
- Térkép: térkép, choropleth, hexbin térkép, kartogram, kapcsolat, buborék térkép;
- Folyamat: akkorddiagram, hálózat, Sankey, ívdiagram, élkötegelés.

Az adatok feltárásának folyamata magában foglalja az adatkészlet felépítésének és összetevőinek, az egyes változók eloszlásainak és a két vagy több változó közötti kapcsolatok vizsgálatát. A feltáró adatok elemzésének célja, hogy elgondolkodtasson adatairól és megalapozza kérdését.

1. Feltáró adatelemzési ellenőrzőlista:
2. A lehető legpontosabban fogalmazza meg kérdését
3. Olvassa el adatait
4. Ellenőrizze a csomagolást
5. Nézze meg az adatok tetejét és alját.
6. Ellenőrizze a számokat, ellenőrizze a számlálást
7. Validálás legalább egy külső adatforrással
8. Készítsen grafikont vagy térképet
9. Először próbáld ki az egyszerű megoldást
10. További kérdések

2. KÉSZSÉG: A STATISZTIKÁK ÉS A VALÓSZÍNŰSÉG MEGÉRTÉSE

Az elején bemutatott azokat az alapvető definíciókat, amelyek szükségesek a leíró statisztika alapjainak és a statisztikailag feldolgozott adatok eredményeinek megértéséhez. Ne feledje, hogy:

1. A medián akkor hasznos, ha egy adathalmazban vannak olyan értékek, amelyek jócskán eltérnek a halmaz számtani átlagától.
2. A hisztogram tökéletesen szemlélteti az adatok változékonyságát egy halmazban.

3. Egy esemény valószínűségi értéke befolyásolja az eredmény várható értékét, és a megközelítés jól működik a portfólió-módszerekben, amikor a kockázatot a sokféleség korlátozza, ahogyan ez a részvények befektetési portfóliójának összeállításakor is láttuk.

A következő egységben megtudtuk, hogy:

1. Az adatsor néhány statisztikája jól leírhatja az adatsor változékonyságát: átlag, medián, szórás.
2. A statisztikák megbízható internetes forrásai nyilvánosan elérhetőek, és az ilyen források nagy előnye az adatok harmonizálása, mint az Eurostat által szolgáltatott adatok esetében is látjuk.
3. A percentilisek hasznosak lehetnek a mérési adatok megértéséhez és értelmezéséhez.

A vizsgálat utolsó egységében két változó korrelációjával foglalkoztunk. Megtudtuk, hogy:

1. A korreláció, más szóval összefüggőség, korrelációs együtthatóval azonosítható és számszerűsíthető.
2. Az értékosztályozás lehetővé teszi a korreláció irányának (pozitív vagy negatív) és erősségének megadását.
3. A példák szemléltetésére használt, általánosan használt táblázatok lehetővé teszik a két változó korrelációs elemzését.

A „Statisztikák és valószínűség megértése” részben bemutatott statisztikai kérdések köre az alapfogalmakra és megközelítésekre korlátozódik. Tudjuk, hogy a statisztikai tanfolyamok korlátozottak a középiskolai oktatási programokban, valamint a bölcsész- és mérnöki tanulmányokban. A statisztika alapjai elengedhetetlenek az „Adatelemzés” modul összefüggéseiben, és egy befejezett „készség”, amely emlékeztethet a megszerzett ismeretekre, valamint ösztönzést és alapot adhat a fejlett statisztikai eszközök elérésére.

3. KÉSZSÉG: HIPOTÉZISEK ÉS ELŐFELTEVÉSEK

Az első lépés egy érvelésben az, hogy tudnia kell, mit próbál bizonyítani és mivel, ezért fontos, hogy ismerje a következtetéseket, a premisszákat és az okokat:

A következtetés mindig egy felvetésre válaszol, ez a felvetés lehet egy kérdés vagy egy vitatható állítás.

Az okok meggyőződések, metaforák és egyéb állítások. Ezek eredendően olyan állítások, amelyek megalapozzák az alkalmazás hitelességének megteremtését. Más szavakkal, ezek magyarázatok vagy igazolások arra, hogy miért kellene hinni egy adott következtetésben.

A következtetést alátámasztó okok az előfeltevések vagy premisszák. Előfeltétel minden olyan állítás, amelyet a szerző a következtetés alátámasztására használ.

A következtetést alátámasztó, nem kifejtett (rejtett) ok a feltételezés. A feltételezés lehet érték (előíró jellegű) vagy leíró.

Az érték vagy előíró feltételezések olyan hitbéli meggyőződések, amelyek arról szólnak, hogy a világnak milyennek kell lennie. A leíró feltevések nem bizonyítottan megfogalmazott vélemények arról, hogy a világ miként alakult, létezik vagy lesz.

Nyilatkozat vagy érveléshez valóban hiteles forrásokra van szükség. Mindig idézd forrásaidat, és mindig támaszkodj megbízható, tájékozott és pártatlan forrásokra.

4. KÉSZSÉG: ÉRVELÉS

A viták kategóriái megalapozzák a koherens érvelés megértését.

A gondolataink rendezésének nagyon hatékony módja az, ha az érvelésünk egyes vitás pontjait kategorizáljuk.

A vitapontok négy kategória egyikébe sorolhatóak: tény, definíció, érték és politika (irányelv)

Az érvelés, mint az érvek bizonyításának módja, sokféle formában fordul elő. A következők hasznos érvelési stratégiák:

- Indokolás általánosítással;
- indokolás analógia útján;
- indokolás okból levezetéssel;
- deduktív érvelés.

A megtévesztés egy érvelési „trükk”, olyan kegyes csalás, amelyet a szerző használhat, miközben megpróbálja meggyőzni Önt a következtetés elfogadásáról.

Három általános megtévesztést különböztetünk meg:

- Olyan érvelés adása, amely téves, vagy téves feltételezéseket igényel, tehát irrelevánssá teszi a következtetést.
- Eltereli a figyelmünket azzal, hogy az információ relevánsnak tűnik a következtetés szempontjából, amikor pedig nem is.
- Olyan következtetés támogatása, amely attól függ, hogy a következtetés igaz-e.

Egyéb gyakori érvelési hibák: *Ad hominem, csúszós lejtő, tökéletes megoldás keresése, egyértelműség, populista hivatkozás, vitatható autoritásra való hivatkozás, érzelmekre való hivatkozás, szalmaember érv, vagy-vagy (hamis dilemma), névvel magyarázás, csillogó általánosság, „vörös hering”, körkörös érvelés, elhamarkodott általánosítás.*

5. KÉSZSÉG: AZ ÉRVELÉS FELÉPÍTÉSE

A legfontosabb témák, amelyekre kitértünk az „adatelemzés epiciklusa” egy olyan háromlépcsős folyamat, amely a következőket tartalmazza:

- Az elvárások beállítása
- Információk (adatok) összegyűjtése, az adatok összehasonlítása az elvárásokkal, és ha az elvárások nem egyeznek,
- Az elvárások felülvizsgálata vagy az adatok rögzítése az adatok és az elvárások egyezésére.

Az érvek megalkotása azt jelenti, hogy bizonyítékokkal és indokkal támasztanak alá konkrét következtetéseket. Ebben a folyamatban a következőkre van szüksége:

- Határozza meg a kérdést, és azonosítsa a lehetséges pozíciókat
- Pontosítsa az alapötleteket érvként
- Fejtse ki részletesebben az érvet
- Vegye figyelembe a kifogásokat
- Keressen tájékozott és pártatlan forrásokat
- Ellenőrizze a forrásokat

Az érvek az indokok előkészítésével, világos és korrekt módon történő megszervezésével kezdődnek. Ebben a folyamatban a következőkre van szüksége:

- Határozza meg az előfeltevéseket és a következtetéseket
- Fejlessze ötleteit természetes sorrendben
- Idézd a forrásokat
- Legyen konkrét és tömör

Az argumentum felépítése során tegye fel a következő kérdéseket:

- Világosan mutatja-e az érvelés, mi is az érvelés következtetése?

- Természetes sorrendben mutatja-e be az ötleteket?
- Megbízhatóak-e az előfeltevések?
- Lehet-e az érvelés világosabb vagy tömörebb?
- Ha igen, mely szavak vagy kifejezések nem egyértelműek?
- Nem keveri-e össze az érvet azzal, hogy egynél több rokonértelmű kifejezést használ ugyanahhoz az ötlethez?

Használja az érvelésében azokat a szavakat, amelyek előfeltevésként és következtetésként szolgálnak:

- Feltétel-mutatók: miatt; mivel; mert; ennek okán; hogy ezt megadta; ez következik abból;
- Következtetési mutatók: ennél fogva; tehát; következésképpen; így; innentől; következésképpen; bebizonyosodik, hogy;

3. MODUL: ADATKOMMUNIKÁCIÓ

Tanulási célok:

- Az eredmények értelmezésének és a megállapítások közlésének képessége,
- az eredmények bemutatásának előkészítése,
- az eredmények vizualizációinak létrehozása (ideértve az adatmesélést és az információs grafikákat),
- alapvető eszközök használata az adatok bemutatásához,
- az eredmények hatékony közlése a célközönség számára (megfelelő tartalommal, stílussal és hozzáállással).

1. KÉSZÉG: ALAPVETŐ ADATKOMMUNIKÁCIÓS KÉSZSÉGEK

Technikai szempontból az **adatkommunikáció** az adat vagy információ átadása az adó és a vevő között: az adó (közlő, forrás) továbbítja az adatokat, és a vevő (fogadó) megkapja. A tanfolyam összefüggésében azonban az adatkommunikációt úgy értjük, mint az adatok elemzéséből és megértéséből származó rálátás átadásának képességét. Az adatkommunikáció alapfogalmainak azonosítása magában foglalja 6 alapkategorizációját: küldő, fogadó, üzenet, adathordozó, protokoll és visszacsatolás.

Számos **módszer** áll rendelkezésre az információk másokkal történő közlésére. A tanfolyam keretein belül bemutatjuk a legfontosabb módszereket, többek között: adatlapokat, jelentéseket, online összefoglalókat és személyes prezentációkat (összefoglalva más módszereket is, például telefonhívásokat, e-mail üzeneteket, fórumokat stb.). Hogyan fogja kiválasztani a megfelelő módszert vagy módszereket a projektjéhez?: Olyan módszerek kombinációját kell választania, amelyek a legnagyobb eséllyel közvetítik az üzenetet. Nagyon fontos, hogy a kiválasztás során vegye figyelembe a **célközönséget** és annak főbb jellemzőit.

Az adatok vizuálása az információk és adatok grafikus ábrázolása. Vizuális elemek, például diagramok, grafikonok és térképek felhasználásával az adatmegjelenítő eszközök közérthető módon biztosítják az adatok trendjeinek, kiugró értékeinek és mintáinak megértését és megértését. A fő teljesítménymutatók a legfontosabb mérőszámok egy adott projektnél. Ez a

legjelentősebb módja a célok elérése szempontjából döntő tényezők értékelésének, amelyek segítenek kiválasztani a legjobb megjelenítési típusokat:

- Milyen történetet akar elmondani?
- Kinek akarja elmondani?
- Meghatározott tendenciákat kíván bemutatni?
- Szeretné bemutatni az adatai összetételét is?
- Össze szeretne hasonlítani két vagy több értékalmazt?
- Az időbeni elrendezettség tényező e?
- Hogyan szeretné megjeleníteni a teljesítés mutatóit?

Az **adatkommunikációs eszközöket** ebben a kontextusban olyan eszközökként értjük, amelyek azokat a vizuális ábrázolásokat segítenek átadni, amelyeket az előzőekben tárgyaltunk, itt: 2. legység (az adott kommunikációs célhoz a legjobb kommunikációs módszer kiválasztása) és a 3. egység (a legjobb vizualizációs típus azonosítása egy adott adat- vagy problémátípus ábrázolására). A legfontosabb eszközök a következők:

- A **táblázatok** számítógépes alkalmazások, melyek adatok szervezésére, elemzésére és tárolására alkalmasak táblázatos formában. A táblázatokat a könyvviteli munkalapok számítógépes analógiájaként fejlesztették ki.
- **Üzleti intelligencia eszközök, beleértve a Dashboard-okat:** Az üzleti intelligencia (business intelligence) szoftver egy olyan alkalmazás, amelynek célja az adatok lekérése, elemzése, átalakítása és jelentések készítése a tájékozott döntéshozás érdekében. Ezen belül a data dashboard egy információkezelő eszköz, amely vizuálisan nyomon követi, elemzi és megjeleníti a fő teljesítménymutatókat (KPI), a méréseket és a kulcsfontosságú adatpontokat, azok figyelemmel kísérése érdekében.
- **Grafikai tervező szoftverek** A grafikai tervező szoftver olyan programra vagy programgyűjteményre utal, amely lehetővé teszi, hogy képeket vagy modelleket vizuálisan manipuláljanak a számítógépen.

2. KÉSZSÉG: FOGALMAZÓ KÉSZSÉG

Egy izgalmas kezdés bevonzza az olvasót, ráveszi, hogy tovább olvasson. Egy jól megírt lead (és cikk) megválaszolja az alábbi kérdéseket vagy legalább is azok többségét. Ez az úgynevezett 5W+1H szerkezet a híreknek: 1. Kiről szól? - Who is it about?, 2. Mi történt (vagy mi az állítás)? - What happened (what's your argument)?, 3. Mikor történt? - When did it take place?, 4. Hol történt? - Where did it take place?, 5. Miért történt? - Why did it happen?, Hogyan történt? -

How did it happen? A kerülendő dolgok magukban foglalják a következőket: szóvirágok, felesleges szavak és sablonos kifejezések.

Az esszében az állítás vagy kijelentés az elsődleges érv, amely az írás legfontosabb szempontja. Néhány tipp, amelyek segítenek a kijelentések és állítások megfogalmazásában és kidolgozásában: **gondoljon a közönségre** (arra törekszik, hogy meggyőzze a hallgatót arról, hogy ötletedet érdemes tovább vinni - vagy hogy egy másik gondolatmenetet már nem érdemes folytatni, kihez szól?); **kutasson** (győződjön meg arról, hogy minden kijelentését bizonyítékokkal támasztja alá); **összpontosítson egy fő gondolatra** (a határozott kijelentések és állítások világosak, tömörök és konkrétak); **ne kezdjen rögtön hozzá** (szánjon időt az átgondolásra, ötletelésre is); **Felülvizsgálat, Felülvizsgálat, Felülvizsgálat** (Soha ne fordítson be első vázlatot!)

Az **esszéekben az érv** általában egy fő gondolat, amelyet gyakran „állításnak” vagy „tézisnyilatkozatnak” hívnak. Bizonyítékok felhasználásával ez az állítás meg is cáfolható. Meggyőzőnek kell lennie. Egy érv megfogalmazásakor javasoljuk, hogy kövesse a következő 3 kezdő lépést:

1. Először meg kell értenie a KÉRDÉST: mit kell tenni? Milyen problémát kell lefedni?
2. Akkor el kell végeznie a KUTATÁST: mit tudunk erről? Mit mondanak más kutatók? Melyek a legfontosabb tények ezekben a vitákban és a vizsgált problémakörben?
3. Ezután térjen vissza a kérdéshez és vizsgálja meg, ön szerint mi a válasz? Ez lesz az ön ÁLLÍTÁSA, azaz álláspontja, amit ezt követően tudományos kontextusban kell majd képviselnie

A tudományos igényű írások esetében az **ellenvetés** egyfajta érv. Mivel minden érvnek kell, hogy legyenek premisszái és következtetései, ezért minden ellenvetésnek is kell, hogy legyenek premisszái és következtetése.

Mindig a legerősebb ellenvetést vegye számításba: Csábító lehet gyenge kifogást választani, amelyet könnyű megcáfolni és megválaszolni. De ha ezt teszi, az megerősíti az ellenérveket, mert nem fog kielégíteni egy szkeptikust. **Egy jó érvelés mindig a lehető legeredményesebb kifogásokat keresi, pontosan és méltányosan mutatja be őket, majd megpróbálja szisztematikusan válaszolni ezekre az ellenvetésekre.**

Mikor teret ad az ellenvetéseknek, akkor ténylegesen elfogadja a másik személy kifogásainak vagy nézeteinek egy részét. Elismeri, hogy talán félig igaza van, vagy hogy jogos az aggodalma. Csak ekkor tudja logikával és/vagy egy megoldással legyőzni ezt az aggodalmat.

Mit kerüljön el: a kulcsszavak túlzott ismételtetését, az érvek közötti csapongást vagy ugyanannak az ismételtetését, a szlenget és az informális beszédet, a gyenge kifogások megcáfolását és az ellenvetések érvek nélküli elutasítását.

3. KÉSZSÉG: SZÓBELI BEMUTATÓI KÉSZSÉGEK

A szóbeli prezentációs készségek azok a készségek, amelyekre szükség van a hatékony és vonzó prezentációk előadására bármely közönség számára. A közönség erőteljes bevonása, aktiválása kulcsfontosságú a sikerhez: ez valós idejű adatok felhasználása ahhoz, hogy személyre szabott és releváns tapasztalatokat és történeteket adjunk át.

A közönség bevonása nélkülözhetetlen a megfelelő hatás eléréséhez. Az előadásának be kell őket vonzania, felhívni a figyelmüket és ösztönözni gondolataikat és megértésüket. Ezt többféle módon, keményen dolgozva el is tudjuk érni:

- A prezentáció tervezése kritikus abból a szempontból, hogy a közönség által értett nyelven és ismereteikre építve adja elő.
- Biztosítani kell azt is, hogy elegendő idő álljon rendelkezésre a kérdések megvitatására és a megbeszélésére.
- Az előadást oly módon kell megtartania, hogy létrejöjjön a kapcsolat a közönséggel.
- A szemkontaktus, a gesztus, a beszélt nyelv és az energia hatékony és lelkes kommunikációt kell, hogy eredményezzen a terem minden szegletével, biztosítva ezzel, hogy a közönség pozitív üzeneteket kapjon rólad és az anyagodról.

A hallgatóság tagjainak aránylag rövid a koncentráló-képességük időtartama, és előadóként keményen kell dolgoznia, hogy folyamatosan megtartsa a figyelmüket. Azáltal, hogy a hallgatóságát kvázi egy utazáson való részvételre csábítja, és folyamatosan tudatja velük, hogy éppen hol járnak, lehetővé teszi, hogy fennmaradjon a figyelmük.

A szemléltető eszközök alkalmazásának legfontosabb előnye az előadás során az, hogy elősegíthetik az információk következetesebb magyarázatát, ami megkönnyíti a bemutatást és a befogadást a hallgatóság számára. Egy beszéd során különböző szemléltető eszközöket használhat: tárgyak, modellek, fényképek, rajzok, grafikonok, diagramok, fóliák, video-, multimédia-prezentációk és akár saját magát is.

Közönsége visszajelzése egy olcsó és hatékony módja annak, hogy fejlessze a nyilvános beszéd tartásának készségét és jobb előadó legyen. Amikor a kritikára gondolunk, ez gyakran

kényelmetlenséget okoz, és tartunk tőle. Ha viszont készen áll a szóbeli prezentációs készségek fejlesztésére, nyitottnak kell lennie a valódi visszajelzések meghallgatására.

A beszéd végkövetkeztetésének szerepe az, hogy jelezze a közönségnek, hogy a beszéd lassan a végéhez közeledik, és segít nekik a beszéd legfontosabb pontjait felidézni. Beszédének lezárása fogja kialakítani a közönség tartós benyomását mindazokról, amiket elmondott. A prezentáció befejezésének módja stílus kérdése (és a stílus személyenként eltérő). Ugyanúgy, ahogyan nincs egyetlen módja annak, hogy elérje a közönségét, nincs egy egyetlen módja annak sem, hogy a prezentáció befejeződjön. Keressen valamit, amely megfelel a személyes stílusának (a forrásgyűjteményben számos példát talál) és találja ki, hogyan alkalmazza azt sikeresen.