\*\*Datasetet rangohen sipas veshtiresise, ku ne baze te veshtiresise qe ata kane, ne fund vleresohen me bonus pike. Bonus piket varijojne nga 0 deri ne 20 pike!

**Kategoria Biznes:**

Telecom\_Customers à 12 pike bonus:

*Te dhenat jane marr nga nje provajder i telefonise dhe internetit, ku jane ofruar shenime per kliente te ndryshem. Disa prej tyre i kane nderprere kontratat me kompanine. Ju duhet te analizoni cilet jane faktoret kryesore qe shtyejne klientet per nderprerje te bashkeveprimit me kompanine dhe te krijohet nje model klasifikues per keta kliente.*

Credit\_Card\_Churn\_Analytics à 8 pike bonus:

*Banka eshte e shqetesuar se pse disa kliente po e nderprejne kontraten per kredi kartele dhe po e leshojne kete produkt. Prandaj ju si shkencetar te te dhenave, keni per detyre identifikimin e klienteve potencial per leshimin e bankes permes nje modeli klasifikues, ashtu qe te ndermerren hapa me kohe ne menyre qe te zvogelohet mundesia e tyre per leshimin e bankes.*

ATM\_Analytics à 15 pike bonus:

*Nje sfide shume e madhe per bankat eshte menaxhimi i ATM. ATM perpos sfidave te sigurise kane edhe sfida logjistike; si mirembajtja dhe mbushja e tyre me te holla ne menyre qe klientet kane gjithmone rastin ti perdorin pa nderprerje ne asnje periudhe kohere. Ne datasetin e ofruar ka transaksione ne ATM qe jane realizuar pergjate nje periudhe te caktuar nje javore. Qellimi eshte te analizohen te dhenat e perdorimit te ATM dhe te nxirren informacione te ndryshme qe permes vizualizimit dhe nderlidhjes se atributeve, mund te ndikojne ne te gjetura kritike per menaxhimin e ATM dhe fondeve te tyre.*

Employee\_HR\_Analytics à 12 pike bonus:

*Largimi i punetoreve nga kompanite eshte problem qe ka impakt ne te gjitha llojet e bizneseve, pa dallim nga pozita gjeografike, industria apo madhesia e kompanise. Largimi i punetoreve ka ndikim te madh ne kosto te kompanise, perfshire ketu punesimi i stafit te ri, trajnimi i stafit te ri. Eshte ne benefit per te gjitha*

*kompanite kuptimi i nxitesve te largimit te punetoreve. Ne datasetin e ofruar ka shenime te punetoreve te nje kompanie, ku disa prej tyre kane leshuar kompanine ndersa disa kane qendruar. Ju duhet te gjeni faktoret kryesore qe ndikojne ne largimin e punetoreve nga kompanite dhe te ndertohet nje model klasifikues mbi kete problem.*

CC\_Analytics 🡪 12 pike bonus:

*Ky dataset permbane kliente te nje banke, te cilet perdorin produktin e kredit karteles. Jane marr atribute nga sjelljes (behaviour) te tyre kunddrejte perdorimit te kredi kartelave per 6 muajt e fundit. Qellimi eshte te analizohen sjelljet e klienteve me kredi kartela dhe te krijohen rekomandime per marketing ndaj klienteve.*

**Kategoria Shendet:**

Stroke\_Analytics à 8 pike bonus:

*Globalisht infarkti eshte shkaktari i dyte i me se shumti vdekjeve ne bote, ku afersisht 11% e vdekjeve vijne nga infarkti. Ne datasetin e ofruar ka shenime per pacient te ndryshem ku disa prej tyre kane pas ose jo infarkt. Qellimi eshte te krijohet nje model klasifikues per infarktin si dhe te analizohet shenimet e atributeve te ndryshme cfar ndikimi kane ne infarkt.*

Suicide\_Analytics à 12 pike bonus:

*Organizata boterore e shendetit - WHO cdo vit mbledhe statistika per vetevrasjet e individeve ne shtete te ndryshme te botes. Qellimi i studimit ne kete dataset eshte identifikimi i faktoreve kyc nga atributet e ofruara se cka ndikon me shume ne vetevrasje dhe te ipen hipoteza sesi mund te zvogelohet ky fenomen.*

**Kategoria Art/Entertainment:**

Movie\_Analytics à 20 pike bonus:

*Dataseti ne vijim ka shenime te mbledhura nga disa platforma (HBO, Netflix, Rottentomatoes).*

*Qellimi kryesor i ketij dataseti gjetja e:*

*- Cfar filma/Seria/Permbajtje ofrohen ne shtete te ndryshme*

*- Grupimi i permbajtjes se filmave ne baze te tekstit te pershkrimit te filmit*

*- Bashkpunimin e aktoreve*

*- Ne vite te ndryshme, si qendron raporti mes Serialeve dhe Filmave, cilat shiqohen me shume*

*- Engine per rekomandim te filmave.*

**Kategoria Edukim:**

PISA\_Analytics à 20 pike bonus:

*Ne fajllin e emertuar "PISA\_selected\_variables\_CY07\_MSU\_STU\_QQQ", keni te dhenat e detajuara*

*nga testi nderkombetar i vleresimit te sistemit te arsimit (testi i PISA-se), per te gjitha shtetet ne te cilat mbahet ky test.*

*Fajlli eshte i gatshem ne dy formate: SPSS dhe .csv. Ne fajllin SPSS, ne "Variable View", e keni*

*codebook te seciles variabel, prej ku mund te kuptoni permbajtjen e emertimit te variables/kolones. Codebook e keni gjithashtu te gatshme ne*

*nje fajll excel, te emertuar "PISA2018\_CODEBOOK".*

*Pyetjet/Variablat te cilat do t'i gjeni ne fajll, ndahen ne kategori kryesore:*

*- Pjesa 1 (Variablat 1-205 jane pyetje te pergjithshme ne te cilat studentet jane pergjigjur rreth sistemit te arsimit, lendeve, ambientit, etj.).*

*- Pjesa 2 (Variablat 206-285 jane pyetjet nga testi per lendet/fushat perkatese nga testi PISA).*

*Vemendje: Databaza mund te reduktohet (variablat ose karakteristikat tjera qe nuk jane me interes) varesisht*

*prej nevojave tuaja per analize, me qellim te zvogelimit te madhesise dhe manipulimit me te shpejte me te dhenat.*

Exam\_Analytics à 0 pike bonus:

*Ne datasetin e ofruar jane shenime mbi nxenes te shkollave fillore dhe progresit te tyre ne disa lende, ju si shkencetar te te dhenave keni per detyre krijimin e nje modeli qe mundet me parashiku se cfar rezultati do te kene nxenesit ne provime. Si dhe analiza te tjera mes nderlidhjeve te atributeve.*