# Information Visualization

# CHECKPOINT I: Visualization Proposal

G14

**1. Domain**

O domínio escolhido foi sobre os utilizadores, os jogos e as diversas categorias de jogos da steam. Escolhemos este tema porque todos os membros do grupo gostam de jogos de computador e além disso também gostávamos de perceber a evolução dos jogos de computador ao longo dos anos, uma vez que esta é a maior plataforma de gestão de direitos digitais. Além disso, achámos interessante perceber como é que os jogos afetam os fatores sociodemográficos de cada país, como por exemplo o número de homicídios, a média de idade, a percentagem de população com depressão e a percentagem da população que utiliza internet.

**2. Dataset**

* + - 1. O grupo vai utilizar o conjunto de dados disponibilizado pela steam no seguinte link: <https://steam.internet.byu.edu/>. E para completar também vai utilizar alguns dados disponibilizados neste site: <https://steamspy.com/> . Para analisar os dados sociodemográficos foram recolhidos dados nos seguintes sites: <https://ourworldindata.org/internet>, relativamente ao número de utilizadores da internet, <https://data.worldbank.org/indicator/SH.STA.SUIC.P5?end=2016&start=2000>, relativamente à taxa de depressão, <https://ourworldindata.org/mental-health#depression>, relativamente à média de idades e o <https://ourworldindata.org/homicides>, relativamente à taxa de homicídios.
      2. **3. Tasks to be supported**

-Tarefa 1: **Explorar** a distribuição dos utilizadores bem como alguns comportamentos/características dos mesmos, tais como o número de amigos que têm e as suas nacionalidades.

-Tarefa 2: **Analisar** a evolução dos jogos e das diversas categorias de jogos e dos seus utilizadores no decorrer dos anos. Ou seja perceber qual foi a evolução relativamente ao número de developers, de publishers, de jogos, de utilizadores e de horas distribuidas pelas diversas categorias de jogo.

-Tarefa 3: **Comparar** os diferentes developers juntamente com os publishers com o número e a qualidade dos jogos lançados.

-Tarefa 4: **Analisar** se o número de horas jogadas em média por cada país tem alguma relação com alguns fatores sociodemográficos, como por exemplo o número de homicídios, a média de idade, a a percentagem de população com depressão e a percentagem da população que utiliza internet.

* + - 1. **4. Example Questions**
* Question 1; Qual a distribuição dos utilizadores e dos seus amigos pelo mundo? (tarefa 1)
* Question 2; Qual é a categoria de jogo com mais jogadores ao longo dos anos? (tarefa 2)
* Question 3; Existe uma relação entre o número de novos utilizadores, o número de novas amizades e o número de jogos lançados no mercado ao longo dos anos? (tarefa 2)
* Question 4; Quais os publishers e os developers que têm jogos com melhor rating? (tarefa 3)
* Question 5; Os publishers e developers que produzem jogos com melhor rating são também aqueles que produzem mais jogos? (tarefa 3)
* Question 6: Será que com o aumento de horas de jogo há um declínio na taxa de homicídio?
* Question 7: Será que nos países que tem maior taxa de depressão joga-se mais horas?
  + - 1. **5. Data Sample**

from table APP\_ID\_INFO

(appid; Title; Type; Price; Release\_Date; Rating; Required\_Age; Is\_Multiplayer)

(63942; ‘Men of War: Vietnam Special Edition Upgrade Pack’, ‘dlc’, 2.99, ‘1970-01-01 00:00:00’, -1, 0, 1)

from FRIENDS

(steamid\_a; steamid\_b; relationship; friends\_since; dataretrieved)

(84679834768374,84397564578332, ‘friend’, ‘2009-12-16 18:51:40’, ‘2013-06-25 08:26:24’, NULL)

from GAMES\_DAILY

(steamid; appid; playtime\_forever; dateretrieved)

(84679834768374, 932048402394, ‘2004-11-15 21:21:39’)

from GAMES\_DEVELOPERS

(appid; Developer)

(932048402394, Valve)

From GAMES\_GENRE

(appid; Genre)

**(**932048402394, Sport)

from GAMES\_PUBLISHERS

(appid; Publisher)

(932048402394; Valve)

from PLAYER\_SUMMARIES

(steamid;lastlogoff; primaryclainid; timecreated; gameid; gameserverip; loccountrycode; loccityid; dateretrieved)

(84679834768374, ‘2004-11-15 16:21:39’, 123456789**,** ‘2006-11-15 21:21:39’**,** 63942, 0.0.0.0, PT130, LIS577, ‘2004-11-15 21:21:39’)

From depression.csv

(Entity,Code,Year,Prevalence - Depressive disorders - Sex: Both - Age: Age-standardized (Percent) (%))

(Afghanistan,AFG,1990,4.42654081684)

from ages.csv

(location; country\_id; data type; database; 1999; 2000; 2001; 2002; 2003; 2004;…)

Afghanistan,"AFG","Population ages 15-64 (% of total)","SP.POP.1564.TO.ZS","55.0646589934074","54.7473132285286","54.5790289355681","54.5008817957377","54.443083552239", …)

From internet\_users

(Entity,Code,Year,Individuals using the Internet (% of population))

(Afghanistan,AFG,2001,0.00472256824217368)