PROJECT 1: PROPOSAL SUMMARY

**Team members:**

Dennis Smirnov

Alexander Singur

Ian Won

Carlos Jaramillo  
 Saumil

**Question:**

Which of these random factors: economic, social or environmental, have the highest correlation with the happiness index?

**Data sets:**

Existing happiness index: (<https://worldhappiness.report/ed/2019/)>

Economy stats:  
 GDP: <https://data.worldbank.org/indicator/ny.gdp.pcap.cd>

Unemployment rate: <https://data.worldbank.org/indicator/SL.UEM.TOTL.ZS>

Tax Revenue: <https://data.oecd.org/tax/tax-revenue.htm>

Ease of doing business: <https://www.doingbusiness.org/en/custom-query>

Non-economy stats:  
> Social:

* Education index ([http://hdr.undp.org/en/data#](http://hdr.undp.org/en/data))
* Alcohol consumption (<https://data.worldbank.org/indicator/SH.ALC.PCAP.LI?year_high_desc=true)>
* Migration data (<https://data.worldbank.org/indicator/SH.ALC.PCAP.LI?year_high_desc=true)>

> Environmental

* Length of day (<https://sunrise-sunset.org/api> / <https://developers.google.com/public-data/docs/canonical/countries_csv)>
* Bio capacity (<https://data.footprintnetwork.org/#/compareCountries?type=BCpc&cn=all&yr=2016)>
* Ecological footprint (<https://data.footprintnetwork.org/#/compareCountries?type=EFCpc&cn=all&yr=2016)>

> Technological

* Internet Access (<https://data.worldbank.org/indicator/IT.NET.USER.ZS>)
* Innovation index (<https://www.globalinnovationindex.org/analysis-indicator)>
* Energy use (<https://data.worldbank.org/indicator/EG.USE.PCAP.KG.OE?view=chart>)