Analysis on how social factors influence the annual increase of HIV/AIDS infection within country level

Apply the linear regression model and test against the model with test set

Wednesday, December 09, 2020

Introduction

Human immunodeficiency virus infection and acquired immunodeficiency syndrome (HIV/AIDS) is a disease caused by the infection of the human immunodeficiency virus (HIV) (Sepkowitz, 2001). If someone is infected, the disease has a long incubation period without any symptom (CDC, 2020), the average incubation period for transfusion cases is about 7.66 years (Anderson & Medley, 1988). As the infection develops, it interferes the functionality of the host's immune system, and increases the risk of being infected by opportunistic infections such as tuberculosis. The symptom after the development of the infection is named as acquired immunodeficiency syndrome (AIDS) (CDC, 2020). HIV spread is caused by unprotected (i.e., no condom) sex (even for oral and anal sex), infectant blood transfusions, hypodermic needle injection (such as taking some drugs), and mother-to-child transmission during pregnancy, delivery, or breastfeeding (Rom & Markowitz, 2015). The virus cannot spread through some body fluids, like sweat, tears and saliva (CDC, 2005). There is no cure or vaccine for this disease, however, the disease can be slowed down by the antiretroviral treatment and with this treatment, victims may have a near-normal life expectancy (CDC, 2020). Currently HIV is not only a form of illness, but also has large impact on the society, such as a source of discrimination and many misconceptions, and also economic conditions (UNAIDS, 2006). And the prevention and remedy of AIDS also has conflicts with religions (McCullom, 2013). This means that AIDS may have some social features and attributes, and according to CATIE, Canada's source for HIV and hepatitis C information, "Many factors in our society, including poverty, physical and sexual abuse, lack of education, homelessness, stigma, addiction, violence, untreated mental health problems, lack of employment opportunities, powerlessness, lack of choice, lack of legal resident status and lack of social support, play a role in HIV infection..." (original text cited from the source) (CATIE, 2018). It is necessary to learn how the society can influence the AIDS spread.

To make the analysis and obtain a tool to predict the AIDS spread, it is worth trying to build a linear regression model which indicates the relationship between the AIDS spread, which is defined as the annual increasing cases within a country level, and several social features as attributes. And in this report, a sample model will be built and tested. The reason to use the linear regression model is that all the social features are quantified and thus all the analysis is made with respect to specific numerical data, and the prediction is also numerical. Linear regression model works based on numerical analysis and makes predictions on continuous change, which is the case for AIDS cases increase per year, based on the change of the social attributes. Meanwhile, the diagnosis of linear regression model suggests if some variables can be rejected for the targeted prediction.

In the rest of the report, the datasets used to make the analysis and test (about the number of new AIDS infections each country per year and the quantified social features in each country per year such as HDI, HAQ), and also the model, how the model is constructed (mathematical notation and variables) and why the model is selected (actually, a linear regression model will be built), the introduction section about the data and the model is referred to as **Methodology**. In the **Results** section, all the figures and tables generated from the predictions made by the model or the diagnosis of model will be displayed, and also may include the analysis on the tables and figures. In the **Discussion** section, all the work done will be summarized, and the meaning of the results will be explained. Also, an analysis of the weakness of the model will be provided and the possible improvements (further steps can be done) will be listed. In the final section (**References**), all

references used in this report are provided.

Results

I got marks deducted form general format in PS3, please provide me some ways to generate clean table with the diagnosis of the model and the dataset from tidyverse and how to number each figure.

References

1. Sepkowitz, K. A. (2001, June 07). AIDS - The First 20 Years: NEJM. Retrieved December 09, 2020, from https://www.nejm.org/doi/full/10.1056/NEJM200106073442306
2. CDC. (2020, November 03). About HIV/AIDS. Retrieved December 09, 2020, from https://www.cdc.gov/hiv/basics/whatishiv.html
3. Anderson, R. M., & Medley, G. F. (1988). Epidemiology of HIV infection and AIDS. Aids, 2. doi: 10.1097/00002030-198800001-00009
4. Rom, W. N., & Markowitz, S. B. (2015). Environmental and Occupational Medicine. Philadelphia, US: Lippincott, Williams & Wilkins.
5. CDC. (2005, February 4). HIV and Its Transmission. Retrieved December 09, 2020, from https://web.archive.org/web/20050204141148/http://www.cdc.gov/HIV/pubs/facts/transmission.htm
6. U. (2006). 2006 report on the global AIDS epidemic. Geneva: UNAIDS.
7. McCullom, R. (2013, February 26). An African Pope Won't Change the Vatican's Views on Condoms and AIDS. Retrieved December 09, 2020, from https://www.theatlantic.com/sexes/archive/2013/02/an-african-pope-wont-change-the-vaticans-views-on-condoms-and-aids/273535/
8. CATIE. (2018). The Social Determinants of Health and Structural Interventions. Retrieved December 09, 2020, from https://www.catie.ca/en/hiv-canada/8/8-1