



---

UNIVERSITY OF WARSAW  
FACULTY OF ECONOMIC SCIENCES

---

**WORKING PAPERS**  
No. 37/2020 (343)

**PREDICTING WELL-BEING BASED ON FEATURES  
VISIBLE FROM SPACE – THE CASE OF WARSAW**

**KRYSTIAN ANDRUSZEK  
PIOTR WÓJCIK**

WARSAW 2020



## Predicting well-being based on features visible from space – the case of Warsaw

Krystian Andruszek<sup>1</sup>, Piotr Wójcik<sup>2\*</sup>

<sup>1</sup> Data Science Lab WNE UW

<sup>2</sup> Faculty of Economic Sciences, Data Science Lab WNE UW, University of Warsaw

\* Corresponding author: [pwojcik@wne.uw.edu.pl](mailto:pwojcik@wne.uw.edu.pl)

---

**Abstract:** In recent years, availability of satellite imagery has grown rapidly. In addition, deep neural networks gained popularity and become widely used in various applications. This article focuses on using innovative deep learning and machine learning methods with combination of data that is describing objects visible from space. High resolution daytime satellite images are used to extract features for particular areas with the use of transfer learning and convolutional neural networks. Then extracted features are used in machine learning models (LASSO and random forest) as predictors of various socio-economic indicators. The analysis is performed on a local level of Warsaw districts. The findings from such approach can be a great help to get almost continuous measurement of the economic well-being, independently of statistical offices.

---

**Keywords:** well-being, economic indicators, Open Street Map, satellite images, Warsaw

**JEL codes:** I31, R12, O18, C14

### Acknowledgements:

Research presented in this article was partially financed by the Polish National Science Center under contract number 2016/21/B/HS4/00670.

## 1. Introduction

GDP per capita or Human Development Index are commonly used measures of economic development on a national and a regional level. The challenge is to measure it on a local level. There are attempts to create synthetic indices based on several dimensions for Poland see e.g. UNDP (2012), Sompolska-Rzechuła (2016), Pomianek (2016), Ciołek (2017).

In recent years, night time lights were commonly used as a proxy of economic activities on regional and local levels (3000 studies since 2000 according to Nordhaus and Chen, 2011). Researchers find a strong positive relationship between night time lights intensity (NTLI) and GDP on a national level, but on a subnational level it is usually weaker and unstable.

One of the possible solutions to measure economic well-being on lower level aggregates like voivodeships, counties or even districts is to use raw satellite imagery. The idea is to extract features from high-resolution satellite images (i.e. features visible from space) and use them to proxy the level of well-being. As relationships between the indicators of well-being and features visible from space are non-linear, machine learning and deep learning tools are often used.

The measurement of economic development, especially in underdeveloped countries with low infrastructure is not trivial. Even for developed countries like Poland, the tool that is able to predict economic indicators based on satellite images would be a great help to get almost continuous measurement of the level of economic well-being, independently of statistical offices and authorities.

The purpose of this article is to analyze the relationship between various socio-economic indicators and features visible from space for the districts of Warsaw. The analysis will be based on high-resolution daytime satellite images gathered from Google maps API and information about points of interest from Open Street Map (OSM). Among various sources of socio-economic indicators we will use real spendings in 2019 (Warsaw budget, 2019), total income, income share in PIT/CIT, population and data about vehicles (Panorama of Warsaw districts, 2017), and also life expectancy (Health condition of Warsaw residents, 2016).

The analysis consists of two parts. In the first part we will apply deep learning and transfer learning to extract from satellite images the features that can be represented as a vector of numbers. In the second stage, machine learning models will be used with the extracted features as explanatory variables explaining each of the socio-economic indicators mentioned

above. Models based on extracted features will be compared with models based only on the data extracted from OSM.

There are three research hypotheses verified in the paper: (1) features visible from space can be used to predict various indicators of socio-economic well-being of Warsaw districts with high accuracy; (2) Features extracted directly from high-resolution daytime images are better predictors of economic well-being than data collected by Open Street Map; (3) Machine learning models that catch non-linearities (for example random forests) offer higher accuracy in predicting the level of economic well-being when compared to a linear regression models, due to highly non-linear relationship between these variables.

The remaining part of the article is structured in the following way: first the importance of the project and actual research in the topic of measuring socio-economic and well-being indicators will be discussed. The second part presents the methods that will be used in the research. The third part consists of data description. At the end the results, final conclusions and further steps will be presented.

## **2. Measuring local well-being from space**

GDP per capita is a common measure of economic development on a national level. However, its usefulness drops when one is interested in assessing well-being at a subnational level. First, in some developing countries regional accounts may not be available or reliable. Second, even in well developed economies regional GDP is a questionable measure. In the calculation of the (regional) GDP per capita the numerator and the denominator can refer to different reference areas. The value of production (numerator) refers to the region in which it has been produced, but it is also the effect of work of people commuting to work from other regions (which should be accounted for in the denominator). This is especially important for large cities and surrounding areas, which are separate administrative units, but create a common labor market. In addition, the use of GDP as a universal measure of well-being is increasingly criticized. Stiglitz et al. (2009) point out that the national accounts measure development with substantial error as the measures of economic production fail to reflect all aspects of human well-being. Although there are attempts to create synthetic indices of the development on a local level based on several dimensions of socio-economic activity – for Poland see e.g. UNDP (2012), Sompolska-Rzechuła (2016), Pomianek (2016), Ciołek (2017), they are rather one-time efforts than a systematic alternative to the mainstream approach.

In UNDP (2012) the authors propose a local human development index (LHDI) as it could be beneficial in case of planning, allocation and supervision of how the European funds are distributed. The components of this indicator are: (1) health, which includes life expectancy and death ratio caused by cancer and heart disease; (2) education as a percentage of children in preschool education and the results from lower secondary school final exams; (3) welfare as average income per capita. LHDI based on listed variables was calculated and grouped for counties (poviats) in 5 categories, where in general the lowest scores were in the Eastern part of Poland.

In research by Sompolska-Rzechuła (2016), the main goal is the assessment of the diversity of social development on a local level. It is done also with the use of local human development index (LHDI) but calculated differently. It consists of indicators related to: health, education, prosperity, health expenditures, educational expenditures and local expenses. The aggregation is on the level of 380 counties (poviats) with the distinction between land (314) and township (66) poviats. The results show that different components were important in these two types of regions. Finally, the LHDI was useful in case of grouping the poviats and finding similarities between them.

Pomianek (2016) aims to find differences and clusters of municipalities (gminas) with the similar level of development taking into consideration rural (1566) and semi-urban (608) municipalities. The socio-economic indicator was based on the variables related to infrastructure, economy and society (15 variables in total). Authors used Hellwig's synthetic measure of the development in order to obtain a single value (in range from 0 to 1) for every of the variables subgroups. Every municipality was classified into one of three classes of the development (high, medium, low). Based on such indicators the differences between municipalities in western and eastern voivodeships were found.

The main goal of Ciołek (2017) was to propose an optimal method of GDP estimation in poviats. Three different approaches were compared: (1) a procedure based on the method proposed by Tokarski (2013), who determined the series of GDP per capita in poviats for the years 2002–2009; (2) GDP was divided in subregions according to the share of individual poviats in the total wage bill; (3) the approach based on the tax revenues of municipalities and counties budgets. It comes out that the third approach was the most accurate and gave the best results.

In contrast to the above mentioned research, this article aims at proposing a comprehensive approach to measure economic development at a subnational level based on continuously available data by applying an innovative method. Differently from traditional

approach, based on imperfect measures of income, or population registers, this project will rely on the satellite images (night-time lights intensity and high-resolution daytime images) supported by machine learning algorithms.

Night-time lights intensity (NTLI) data is recently used as a rough proxy of economic activity on a regional and local level (3000 studies since 2000 according to Nordhaus and Chen, 2011). NTLI is based on satellite images collected by sensors installed on satellites and is increasingly used to monitor human impact on Earth's surface. Numerous studies have been carried out for verification of the relationship between NTLI and population, economic wealth (GDP), urbanization, electricity consumption on national and subnational levels (e.g. Bennett and Smith, 2017; Jasiński, 2017). However, in most cases only linear or log-linear relationships were verified. Strong correlations between NTLI and GDP are found at the country level (e.g. Henderson et al., 2012, Pinkovski and Sala-i-Martin, 2016). However, Bickenbach et al. (2016) show that the relationship between the growth of NTLI intensity and regional GDP growth is unstable on subnational level. Within countries NTLI seems to be more strongly correlated with population density than welfare. The relationship between nighttime lights and remunerations or other measures of income at the local level appears weak and non-linear (Gennaioli, 2014; Mellander et al., 2015).

There are important shortcomings of the NTLI data – limiting the DMSP-OLS data on the 0-63 scale does not allow to correctly distinguish regions with large income differences (for example the core metropolitan areas of New York City and Mexico City are top-coded, although there are large income differences between them – see Lessmann and Seidel, 2017). What is more the NTLI data are not directly comparable over time due to independent scaling every year, even if just one type of NTLI data is used (e.g. DMSP-OLS). Therefore, usually the studies concerning the relationship between NTLI and wealth are time invariant and concentrate on a single year (e.g. Elvidge et al., 2009 analyzes data for the year 2004 and Elvidge et al., 2012 just for 2006).

Another innovative solution to overcome the limitations of nighttime lights intensity data, especially its relatively low spatial resolution, is to analyze high-resolution daytime satellite images collected for example from Google Static Maps API. Powerful image recognition tools (convolutional neural networks, CNN) are used to extract meaningful features from images. They can be used for example to derive the number and density of buildings, the prevalence of shadow area as a proxy for building height, the number of cars, density and length of roads, type of farmland, roof material, etc. Therefore, for each area the set of satellite images is converted into a meaningful set of characteristics which can be used in a statistical model as

predictors for an indicator of economic well-being. Due to highly non-linear relationships often artificial neural networks or random forests are used in this case. As the number of extracted features can be larger than the number of areas analyzed, LASSO regression can be applied to identify the most important predictors (e.g. poverty rates and average log consumption in Engstrom et al., 2017). Such a model can be then applied to accurately predict the level of economic well-being even on small areas.

The disadvantage of using high-resolution daytime images is the lack of long series of historical data. However, in the near future such data might be available with even daily frequency. With a model accurately predicting the level of economic well-being based on satellite images, good proxies for regional and local level of economic development might be available much before the official statistical data are collected, processed and published.

Jean et al. (2016) used survey and satellite images from African countries such as: Nigeria, Tanzania, Uganda, Malawi and Rwanda to predict households consumption expenditures as well as asset wealth. In order to gather image features, CNNs and transfer learning techniques were used. In addition, performance of this approach was compared with the results obtained from prediction based on night-time lights. It comes out that using satellite imagery gives higher accuracy. Although  $R^2$  varies a lot between countries, dependent variables and different poverty measures, it was up to 75% which is a promising outcome.

Most research is conducted in developing countries where estimation of poverty is crucial for policy makers. Tingzon et al. (2019) used Demographic and Health Survey in which social, economic and health related informations were collected. Satellite images were taken from Google Static Maps API. In addition, the Open street map (OSM) data, which has information about the number of buildings, roads, and other points of interest in given area, were used. The Deep Learning techniques were applied to obtain features from images including Transfer Learning, with the use of VGG16 architecture. On top of that random forest regression models were performed to get the final prediction. The best model achieved 0.63  $R^2$  for asset-based wealth.

Babenko et al. (2017) try to predict poverty for Mexico. In contrast to research mentioned before, the authors used Planet and Digital Globe imagery instead of Google Maps. Two different CNN architectures (VGG, GoogleNet) were checked and the poverty benchmark was a combination of household and income per adult equivalent surveys. The final model on urban and rural areas obtained  $R^2$  of 57% and 64% when applied on urban areas only. The

solution based on municipalities data shows that it is possible to obtain an end-to-end solution with satisfactory estimates of poverty.

Chen et al. (2016) used a different approach. The goal of this research was to improve the mapping of housing rents in Guangzhou, China. It was done with the use of open-source data at the level of neighborhood committee (NC). The accurate information about housing rent is important in case of monitoring and evaluation of local residential markets. It can be also used as a detector of housing markets inflation by measuring price-to-rent ratio. It is also mentioned to be useful in case of designing new urban development and renewal of declining urban areas. The main difference with previous research is the fact that instead of satellite imagery, the points of interest (POIs) from OSM were used. Therefore, CNNs were not applied and only ensemble learning in order to predict housing rent was used. This research shows that usage of POIs and some additional features (the nighttime lights intensity data from NOAA/NGDC and Normalized Difference Vegetation Index, NDVI from Landsat image) can help in estimation of rents. Most importantly, this data is updated much more frequently than conventional statistical data.

The above literature review allowed to identify an important research gap – the lack of comprehensive studies using satellite data to predict economic well-being on a local level. Previous studies often analyze the relationship between NTLI and economic well-being using simple statistical tools: linear correlation or linear regression, while the actual relationships are non-linear and often different for urban and other areas (Wójcik, 2018b).

There is a lack of analysis based on satellite imagery on a local level for Poland. This research is probably one of the first approaches to predict socio-economic and well-being indicators for Parts of Poland with the use of innovative methods such as CNNs and machine learning. The article aims to answer the questions whether features visible from space can be used to accurately predict such indicators. The approach based on satellite imagery will be compared with the approach based on Open Street Map data (POIs) only. In addition, LASSO and random forest models will be compared to check whether algorithms that can catch non-linearities are better in case of prediction accuracy.

### **3. Methodology**

Our empirical approach for predicting economic indicators consists of two parts. The first part is based on Convolutional Neural Networks and transfer learning and its goal is to extract features visible from space from high-resolution daytime satellite images. The second part consists of application of predictive models based on those features where random forest will

be compared with LASSO. OSM data and features extracted from satellite images will be used as model inputs to predict various socio-economic indicators on the level of Warsaw districts. Alternatively, models with only OSM data only will be used for comparison.

### 3.1. *Cross-Validation*

Cross-validation is a popular technique used for model estimation, comparison and selection of best hyperparameters. The basic concept is a hold-out method in which, data is split into train and test set. In such way, model is built on training set and tested on the test set.

The simplest option of cross-validation is k-fold random subsampling. In this technique, the hold-out method is repeated k-times, so k-pairs of train and test set are generated. For each combination of train-test set, model is trained and tested. The final performance of the model is an averaged prediction of all k-sets.

There is a specific cross-validation method when  $k = n$ . It is a leave-one-out cross-validation in which each individual observation serves as test set. The train set consists of remaining  $n - 1$  observations.

In general resampling methods are used to avoid biased estimates. In addition, benefit of cross-validation techniques is overfitting prevention and is widely used in hyperparameter selection process (Berrar, 2018).

### 3.2. *Lasso*

LASSO (Least Absolute Shrinkage Selector Operator – Tibshirani, 1996) is a popular machine learning algorithm used for the selection of the most important predictors in the model. This algorithm is a specific example of regularization (also called shrinkage). The general idea is to estimate a regression model with an additional constraint, which causes the estimated parameters to be closer to 0 (shrinking to 0) or some of them even be equal to 0. The constraint requires that the sum of absolute values of all model parameters (apart from the intercept) is as small as possible. Adding a penalty in the optimization results in searching for parameters that fit the data well, but are as small as possible (nearest 0). It allows to decrease the impact of non-relevant features or even remove them from the model. The additional constraint is added to the optimization problem (e.g. for linear regression minimizing residual sum of squares) with a positive weight (hyperparameter lambda), which optimal value should be found with the help of cross-validation.

### 3.3. *Random forests*

Decision and regression trees are popular tools used in the regression tasks (continuous dependent variable) and in the classification tasks (discrete dependent variable). A tree is constructed by dividing a data set into two or more subsets of observations with respect to the values of one of the explanatory variables (predictors). Each subset is then further subdivided according to the same algorithm the process is repeated for subsequent subsets until some stopping criterion is adopted. The attractiveness of decision trees is due to the fact that they give easy-to-interpret results. These results can be directly translated into decision rules and applied to a data set. However, decision trees are characterized by a worse prediction compared to more advanced tools. One of the important disadvantages of the trees is the problem of overfitting – fitting too much to relationships observed in the training sample which do not necessarily repeat in the new data (test sample). One of the solutions of the problem of overfitting is bagging (bootstrap averaging). It consists in multiple estimation of the same model on different subsamples. Subsamples are created using random sampling with replacement of observations from the original set (bootstrap samples). On each subsample the same type of model is estimated and predictions are generated from each obtained model. The final prediction from the model is created by combining the results of all classifiers. One of the specific examples of bagging applied to tree models is a random forest – a collection of trees estimated on bootstrap samples. In addition, during the construction of a single tree at each division only a random subsample of predictors is considered from the full set of all predictors, which introduces additional randomness and decreases correlation between models. Combining predictions from many (weakly correlated) models trained on different subsamples reduces the problem of overfitting. Training a random forest requires selection of two model hyperparameters – the number of trees in the forest and the number of predictors randomly selected at every division.

### 3.4. *Neural Networks*

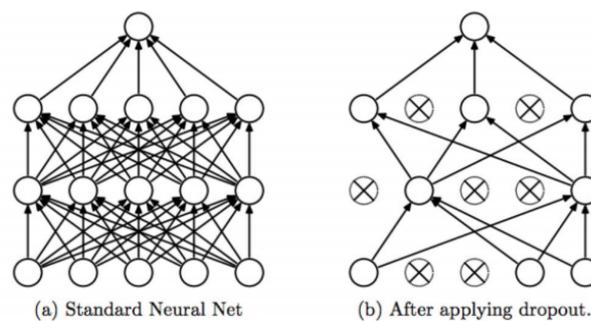
Artificial Neural Networks (ANN) are a type of computer software inspired by biological neurons. The human brain is able to solve difficult problems, but each neuron is responsible only for a very small part of it. Similarly, neural networks are made up of cells that work together to get the result.

Feed forward neural network is one of the first and the simplest architectures. It consists of three types of layers. The input layer, which is responsible for entering data into the network. The next one is hidden layer, which number can vary, so networks are divided into single or multilayer, based on the number of hidden layers. The last layer is called the output layer and is used to determine the output values of the network. The deeper networks can solve non-linear and much broader class of problems, where single or layers networks can be applied only to relatively easy tasks.

Each neuron has an activation function which is a function, on the basis of neuron output value is calculated. There are many activation functions e.g. sigmoid, softmax or rectified linear unit (ReLU). They have various practical meanings and their choice may also determine general characteristics of the network.

ANN consists of dense layers, in which every neuron is connected with every neuron in the next layer. Often the dropout parameter is used which is responsible for “dropping” random neurons when training the network. It is responsible for regularization and prevents overfitting. During the training process only some of the neurons of the layer remain, and the rest are ignored.

**Figure 1. Sample fully connected neural network with and without dropout**



Source: Li, Johnson & Yeung, (2017).

Backward propagation of errors is a method used in supervised neural networks. Algorithm given the final loss value of the neural network, works from the top layers to the bottom layers, calculating the contribution of each parameter in the loss value. The algorithm makes corrections of weights based on gradient optimization methods. The correction of the network weight vector is based on minimization error measure function.

### 3.5. Convolutional Neural Networks

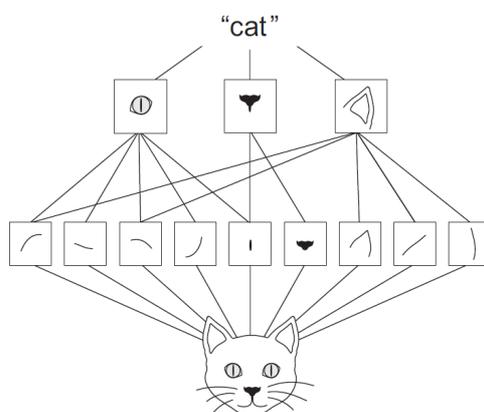
CNN's are based on Feed Forward Neural networks. They are a result of research on the visual cortex and used since the 1980s, mainly applied in image classification and image and video recognition. In recent years CNNs were able to achieve results that highly exceed human capabilities (Geron, 2018, p. 351-352). Its name comes from a mathematical operation called convolution, and CNNs are simply neural networks where at least one of the layers uses convolution instead of matrix multiplication (Goodfellow et al., 2016).

The main difference between a dense and convolutional layer, is that neurons are not connected with every pixel of image. It causes that dense layers learn global patterns and convolutional – local patterns. This characteristic makes CNNs more suitable in image analysis, especially when input feature map is big, e.g. a image with 100x100 pixels size and the first layer consists of 1000 neurons, it would produce in the first dense layer 10 million connections (Geron, 2018, p. 353).

As CNN learns local patterns, it gives important properties (Chollet, 2018):

- 1) discovered patterns are translation invariant. It means that the network is capable of identifying the same pattern in different parts of images. Let's assume that the task is to recognize whether a particular object (e.g. cat) is in the picture. In such scenario one does not care where the cat is, but if the cat is in the image. In case of ANN, the network will learn a specific place of the object from every input image, which is inefficient and would end with poor performance.
- 2) CNN learns hierarchical structure. It means that first layer of the network learns general characteristics and in the next layers combines information from previous steps to obtain more complex characteristics (as presented in the Figure 2).

**Figure 2. CNN features extraction example**



Source: Chollet (2018).

One of the problems in computer vision is overfitting, which is caused by too few samples. In such a situation, the model is unable to generalize to new data. Data augmentation is a popular technique which is used to artificially expand and increase the diversity of the dataset without collecting new observations. The popular augmentations are: horizontal and vertical flips, rotations, random crops etc.

### 3.6. *VGG and ResNet*

VGG is a convolutional neural network proposed by Simonyan and Zisserman (2015). Authors presented few different specifications and at that time CNNs with 16 and 19 layers were one of the biggest architectures. In 2014 ImageNet Large-Scale Visual Recognition Challenge (ILSVRC - contest that evaluates algorithms for object detection and image classification) these solutions achieved first place in localization and second in classifications tracks.

In 2015 Residual Neural Network (ResNet) was presented (He K. et al., 2015). This solution won the ILSVRC 2015 competition with a top-5 error rate of 3.57%, which beats human level performance. Both architectures are widely used in different object localization and image classification tasks.

### 3.7. *Computer Vision*

Computer vision has broad applications. From the simplest as image classification through classification with localization to object detection. In this article, semantic segmentation approach will be used instead of instance segmentation as the overall area of a given class will be measured and single objects area will not be taken into consideration.

The objective of semantic segmentation is to classify each and every pixel of the image to one of the predefined classes. The output of such an approach is a matrix with values corresponding to the class objects. In comparison with the object-detection approach, there are no bounding-boxes around the objects and single instances of a given class are not distinguished. There are various applications of semantic segmentation approach. It is used in medical images analysis, e.g. in tumor diagnosis as well as in autonomous vehicles which requires very precise classification of every pixel. In recent years it has also become a popular approach in satellite images analysis.

Probably the most popular task in computer vision is image classification. In such scenario architecture consists of few convolutional and max pooling layers which are responsible for detecting characteristics. It detects characteristics from most basic to complex

ones in the further layers. After that step, there is a classification phase which takes extracted features as inputs and outputs one of predefined classes. This part consists of dense layers.

In semantic segmentation, classification part is not needed as every pixel is classified and it is done with decoder-encoder type of architecture. It consists of two parts: 1) encoder is responsible for downsampling the input size and storing information about the objects, their shape and size; 2) decoder takes this information and creates the segmentation map.

In this article, the U-Net architecture will be used (Ronneberger et al., 2015). The main benefit of this approach is that it doesn't need a lot of training samples and relies on data augmentation. The architecture consists of contracting and expanding paths. Contracting is responsible for capturing the context and expanding for precise localization. Architecture consists of 3x3 convolutional layers with rectified linear unit (ReLU) activation function and 2x2 max-pooling operations for downsampling on contracting path. For expanding path, max-pooling operations are replaced with up-convolutions and are combined with concatenations with high-resolution features from contracting path. The crop operation is needed as in every up-convolution operation, pixels from the border are lost.

### 3.7. *Transfer Learning*

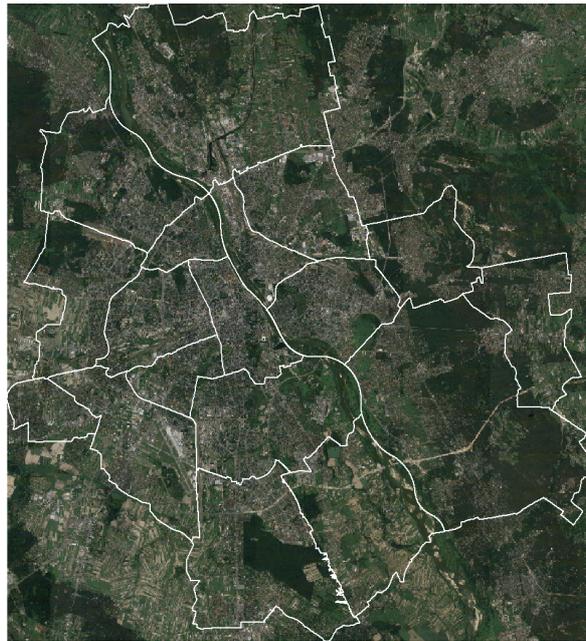
Estimation of CNN capable of extracting features from thousands or millions of images with good quality is a computationally demanding process. In order to solve such problems easier the transfer learning approach is used. The main idea behind it is to take a neural network initially pre trained on another similar dataset and only fine-tune the weights or even just train the last layers of the whole architecture for a given task. Transfer learning is usually used where one does not have enough data to obtain satisfying results. Most of the time the ImageNet, which is collection of many categorized images, is used as a base dataset on which popular architectures such as VGG, ResNet or InceptionNet are pre trained. In case of semantic segmentation and U-Net architecture, the transfer learning can be used as well. The only difference is that the popular architectures (backbones) such as VGG, ResNet, InceptionNet can be used only in the encoder part. It is still very useful, but the decoder part needs to be trained from scratch.

## 4. Data

### 4.1. Satellite imagery

The source of imagery data is Google Maps API. The data for the whole Warsaw with districts distinction (18 districts) was collected in August 2019 – see Figure 3 for a general overview.

**Figure 3: Combined satellite images overlaid with the borders of Warsaw districts**



Overall, 6318 high-resolution daytime images of 1024 by 1024 pixels were collected, which corresponds to over 1.5 GB of data. Images are of Google zoom size 16 which corresponds to around one square kilometer. The coordinates for which the data was downloaded are: longitude 20.85169 to 21.27115, latitude 52.09785 to 52.36815.

The Open Street Map data was used for localization of buildings, streets, rivers and lakes, green areas, public transport stops, bike rental stations, fuel stations, supermarkets and malls. These data<sup>1</sup> was used as: 1) model features, for which data was aggregated to district level and following characteristics were calculated: buildings number and area, number of public transport (bus, tram, train, subway) stops, number of bikeshare and fuel stations, length of roads as well as area of green and water and number of malls and supermarkets; 2) input for labeling images. For this purpose, precise matching of images and POIs localization was

---

<sup>1</sup> To collect the data, R package - osmdata was used. Objects that can be identified in satellite images, and on the other hand, those that may be related to the wealth of a given area were selected. Information about the height of buildings or the type of roof on buildings are also analyzed (Engstrom et al., 2017), but there was no such data available in this project, and manual labeling would be too time-consuming.

needed. Unfortunately, the pictures were taken from a slight angle and some of the objects visible from space were not marked in OSM data, which negatively affects the quality of the model. Another challenge was connected with green and water areas as lot of them were not marked at all in OSM (see Figure A3 in appendix). To address this problem, a different way of extracting was applied.

In Table 1., OSM data is summarized with the mean and a coefficient variation (CV) calculated for every variable. The biggest CV value is observed for the length of trunk roads and it is caused by the fact that in some of the districts there are no such roads. On the other hand, the lowest variety is for the area of buildings and a number of bus stops. In general, the CV is in range from 45% to 153%.

It is also visible in Table 1, that some of the variables in OSM are highly unreliable. The problem with green areas is mostly visible in Wesoła district where the value is drastically low, while Wesoła is one of the greenest districts in Warsaw. What is more, the total number of subway stops is equal to 63 (there are 34 subway stops<sup>2</sup>). This is caused by the fact that in some stations more different entrances to subway are marked as separate subway stops. Such inaccuracies could negatively impact model accuracy.

#### 4.2. *Economic indicators*

The economic indicators data comes from several sources. The spendings were collected from the 2019 Warsaw Budget. The information about total income, income share in PIT/CIT, population (total, men, women) and vehicles (total, passenger) were gathered from Panorama of Warsaw districts (2019). In addition, life expectancy from the report “Health condition of Warsaw residents” (2016) was also used. The data does not cover the same year, but it should not affect the results strongly. One can assume that changes of these indicators at the level of districts are rather small from year to year.

In Table 2., the values of economic indicators are presented together with the mean and coefficient of variation across districts. CV ranges from 3% for life expectancy up to 101% for income variable. It shows that prediction of life expectancy should be rather an easy task as the values are very similar among the districts. In addition, it is visible that features describing population as well as vehicles are collinear between different specifications of these characteristics.

---

<sup>2</sup> [https://pl.wikipedia.org/wiki/Lista\\_stacji\\_metra\\_w\\_Warszawie](https://pl.wikipedia.org/wiki/Lista_stacji_metra_w_Warszawie)

Table 1. Values of features obtained from OSM for Warsaw district

District	area_buildings	num_buildings	pubtrans_bus	pubtrans_train	pubtrans_tram	pubtrans_subway	num_bikeshare	num_fuel_stations	str_length_primary	str_length_residential	str_length_secondary	str_length_tertiary	str_length_unclassified	str_length_trunk	area_water	area_green	number_mall	number_supermarket
PRAGA-POLUDNIE	3008345	7473	319	19	52	0	28	102	14051	110769	29305	43591	2192	0	853209	902701	6	14
BIAŁOLEKA	4355817	12435	432	7	32	0	10	87	15186	242483	17297	70730	12498	0	449764	8606595	6	18
MOKOTÓW	4265839	9363	421	3	66	10	39	125	22264	139738	66650	45299	3806	0	950259	3548119	3	7
BIELANY	2225264	5904	316	1	57	9	16	62	15082	85958	28519	37687	3346	0	455280	1974975	2	13
OCHOTA	1774250	2614	133	13	44	0	27	34	11011	47539	23105	11242	622	0	19911	502269	2	4
BEMOWO	2070012	5324	167	0	59	0	17	58	0	100161	31939	27476	2453	9733	53473	845219	8	11
REMBERTÓW	1123002	4836	165	3	0	0	7	4	9196	70369	5176	19014	2212	0	26670	408214	0	4
TARGÓWEK	2599218	6795	261	2	21	2	18	99	1740	102783	24287	38226	2815	10159	47688	1391163	7	8
WAWER	4201548	18739	459	20	0	0	1	41	18381	303046	40708	60280	4755	0	1934303	7087154	2	16
URSYNÓW	3385563	8916	285	6	0	10	23	52	17159	114034	30592	36571	4246	9675	260536	3783516	3	15
ŚRÓDMIEŚCIE	3093402	4647	327	11	73	16	82	28	18668	61882	32041	46364	1564	0	1010281	430033	7	3
WESOŁA	1123001	5058	78	4	0	0	0	16	7622	98516	5765	13847	1976	0	33602	24723	0	5
URSUS	1274779	3853	147	4	0	0	5	10	0	57777	8158	20895	575	3697	15012	142577	1	7
PRAGA-PÓLNOC	1368982	2083	147	8	62	5	27	18	2804	32414	27898	16032	1139	2806	1241328	628274	2	2
WOLA	2687147	5062	270	18	82	7	33	133	9067	70290	37773	26538	2049	268	21936	781831	3	5
WILANÓW	1652998	4737	120	0	0	0	8	6	0	88153	21549	18346	2489	0	1270335	8188930	0	5
WŁOCHY	2767635	7545	185	11	14	0	8	94	13426	108979	30849	23537	1095	20117	126583	975640	4	7
ŻOLIBORZ	974107	2636	150	1	33	4	12	22	9488	32441	19246	16546	2562	9712	534184	474750	2	3
mean	2441717	6557	243	7	33	4	20	55	10286	103741	26714	31790	2911	3676	516909	2260927	3	8
CV	45%	59%	47%	89%	86%	136%	92%	74%	67%	64%	52%	51%	89%	153%	107%	122%	77%	60%

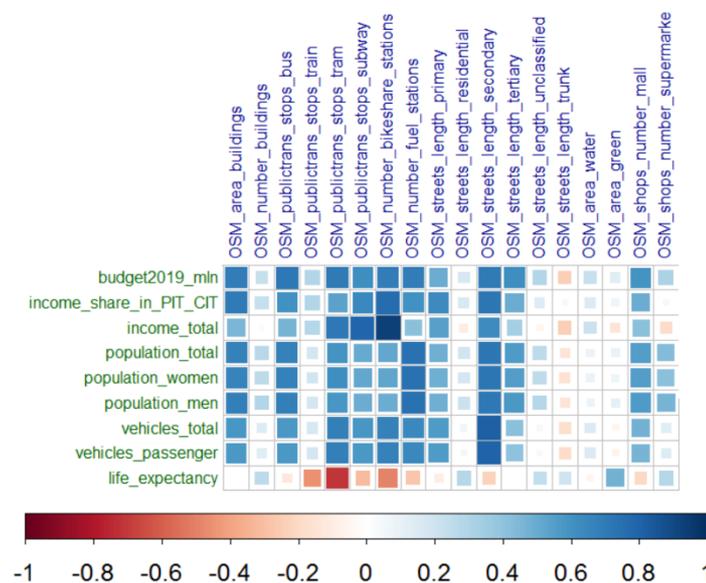
Table 2. Values of socio-economic indicators of Warsaw district

District	budget2019_mln	income_share_in_PIT_CIT	income_total	population_total	population_women	population_men	vehicles_total	vehicles_passenger	life_expectancy
PRAGA-PÓLUDNIE	728	39680	173019	178726	97724	81002	156727	136596	75
BIAŁOLEKA	620	48231	89736	119374	62915	56459	63344	51562	77
MOKOTÓW	855	100767	371405	217577	119525	98052	263881	217997	76
BIELANY	539	20244	105954	132026	71902	60124	103997	89174	75
OCHOTA	386	32836	127462	83081	45436	37645	90946	76894	74
BEMOWO	429	24707	57179	122210	65191	57019	103959	84701	78
REMBERTÓW	120	8427	16471	24148	12603	11545	19323	15350	78
TARGÓWEK	555	35061	93933	123941	66892	57049	82058	66511	75
WAWER	371	27944	47702	75991	40149	35842	52590	42205	77
URSYNÓW	572	41908	101832	150273	80125	70148	88412	73853	80
ŚRÓDMIEŚCIE	728	96723	503254	117005	64476	52529	152856	125140	74
WESOŁA	123	6795	8183	25106	13034	12072	14798	11743	78
URSUS	288	16479	28604	59261	31527	27734	34850	29376	76
PRAGA-PÓLNOĆ	443	14807	101684	64904	34924	29980	46918	39720	71
WOLA	689	59687	261374	140111	76819	63292	109758	91003	73
WILANÓW	271	16270	23655	37511	19930	17581	20814	16548	81
WŁOCHY	225	52876	77508	41929	22313	19616	50448	40523	77
ŻOLIBORZ	263	10557	74029	51441	28353	23088	45189	36518	77
mean	456	36333	125721	98034	52991	45043	83382	69190	76
CV	46%	74%	101%	55%	55%	54%	72%	73%	3%

## 5. Results

To initially assess which features will be good predictors of individual economic indicators, correlation between OSM data and the indicators were performed. It comes out that the highest correlation is between the number of bikeshare stations and total income. As most of the bike rental stations are in the center of the city, it explains this result. It is also worth to mention that 3 out of 6 features of streets length (different types of streets) are relatively uncorrelated with all the economic indicators. In addition, green areas, which have poor coverage and area of water are not highly correlated with any of the dependent variables.

**Figure 4: Correlation plot between OSM features and economic indicators**



To predict economic indicators two different models were used: 1) LASSO, which is a linear model with a regularization part that has the possibility of excluding least important features; 2) random forest, which was selected because it is able to find non-linearities in the data. We use two approaches: 1) models based only on OSM data; 2) models based on OSM and features extracted from semantic segmentation approach. Due to the small number of observations (18 districts) the leave-one-out cross validation method will be applied and both models will be estimated for each of 9 economic indicators as a dependent variable. The quality of predictions will be measured and compared with Mean Absolute Percentage Error (MAPE). In addition feature importance will be analyzed to check which features are most important. In the second approach, with extracted features, it also will be beneficial to check how important those features are. Such approach allows to verify all research hypotheses.

### 5.1. *Models with OSM features*

First, the LASSO model based on OSM features only was estimated for each of the economic indicators. The summary of its results in terms of the ranking of variable importance is presented in Table 1 (the lower the number, the more important the variable is). It appears that in general the most important feature is the number of bikeshare stations and the number of fuel stations. Both of them were important in 7 out of 9 models and in 6 models were in top 5. On the other hand, area of buildings, area of greens and areas of water are the least important features.

In LASSO model, the lambda parameter was selected individually for every model. Despite that, the prediction error is high, see Table 3. As it was expected, the lowest MAPE is for life expectancy model and the highest for income. In case of life expectancy which is used in LHDI, most important features are connected with access to public transportation, malls and supermarkets. Unintuitive is the fact that close access to green and water areas is not so important, but it can be related to quality of these features in OSM. The other important model is income share in PIT/CIT, its often used as the measure of income. The MAPE is equal 60% and interesting is the fact that only three variables are important: number of bikeshare stations, length of secondary roads and area of buildings. Worth mentioning is the fact that in case of life expectancy, area of buildings is one of the least important features. In the approach based on OSM and satellite images data, this variable will be calculated based on CNNs.

Random forest gives very similar results to LASSO. In 6 models, the number of bikeshare stations is the most important feature, see Table 4. Next most important features in general are number of fuel stations, and public transport features. In this model the area of green places is one of the least important features similarly as in the LASSO model.

Life expectancy model in comparison to LASSO, achieved slightly better accuracy. The big difference is in importance of the features. Number of bikeshare stations is most important and surprisingly area of water is third. From public transportation features, number of train and tram stops look promising. In comparison to LASSO, number of malls and supermarkets are placed on seventeenth and twelfth correspondingly. In case of income share in PIT/CIT, once again bikeshare stations, area of buildings and street length were ones of the most important features, when area of green and water placed sixteenth and twelfth. MAPE for this model is slightly bigger (60.5% in comparison to 60% for LASSO).

Table 3. Ranking of feature importance and error measures for LASSO models based on OSM features only

variable	budget2019_mln	income_share_in_PIT_CIT	income_total	population_total	population_women	population_men	vehicles_total	vehicles_passenger	life_expectancy
OSM_area_buildings		3	13	14					14
OSM_area_green				13			8	8	16
OSM_area_water			12	15					15
OSM_number_bikeshare_stations		1	2	6	4		2	2	5
OSM_number_buildings									9
OSM_number_fuel_stations			4	5	1	1	4	4	8
OSM_publictrans_stops_bus				7	3	2			7
OSM_publictrans_stops_subway			3						6
OSM_publictrans_stops_train			5	4			3	3	2
OSM_publictrans_stops_tram				3	2				4
OSM_shops_number_mall			6	2			1	1	1
OSM_shops_number_supermarket			1	1					3
OSM_streets_length_primary			7	12			7	7	10
OSM_streets_length_residential				11					
OSM_streets_length_secondary		2	11	10	5	3	5	5	12
OSM_streets_length_tertiary			10						
OSM_streets_length_trunk			9	9			6	6	13
OSM_streets_length_unclassified			8	8					11
MAPE	64,9%	60,0%	84,6%	53,9%	63,2%	63,6%	55,0%	58,8%	6,0%

Table 4. Ranking of feature importance and error measures for Random Forest models based on OSM features only

variable	budget2019_mln	income_share_in_PIT_CIT	income_total	population_total	population_women	population_men	vehicles_total	vehicles_passenger	life_expectancy
OSM_area_buildings	7	7	11	7	7	7	9	9	14
OSM_area_green	12	16	17	15	15	15	15	16	10
OSM_area_water	9	12	10	13	12	13	11	11	3
OSM_number_bikeshare_stations	1	3	2	1	1	1	2	1	1
OSM_number_buildings	11	11	12	6	8	5	10	10	5
OSM_number_fuel_stations	5	10	9	5	4	6	5	5	11
OSM_publictrans_stops_bus	4	9	8	4	5	4	6	6	13
OSM_publictrans_stops_subway	14	5	4	16	16	16	12	12	16
OSM_publictrans_stops_train	17	13	13	17	17	17	16	15	2
OSM_publictrans_stops_tram	6	2	3	10	10	12	4	4	4
OSM_shops_number_mall	3	4	6	3	3	3	8	7	17
OSM_shops_number_supermarket	15	17	14	11	11	10	17	17	12
OSM_streets_length_primary	8	1	1	9	9	9	3	3	8
OSM_streets_length_residential	16	14	16	14	14	14	14	13	6
OSM_streets_length_secondary	10	6	5	8	6	8	1	2	9
OSM_streets_length_tertiary	2	8	7	2	2	2	7	8	15
OSM_streets_length_trunk	18	18	18	18	18	18	18	18	18
OSM_streets_length_unclassified	13	15	15	12	13	11	13	14	7
MAPE	39,0%	60,5%	135,2%	51,8%	53,2%	50,1%	67,8%	71,1%	2,2%

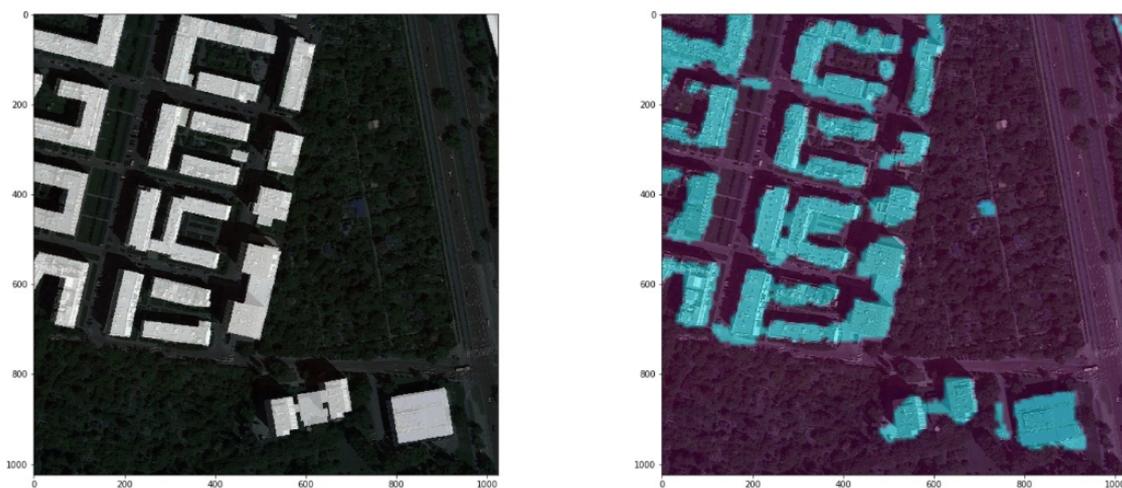
In both Lasso and Random Forest, the smallest error is in the life expectancy model, but it is due to small variance across districts. The population models look quite promising. The worst result was achieved by the total income model. In general, for both: LASSO and random forest, fitness of models is not as good as expected, which is probably caused by small amount of data.

### 5.2. Model with OSM and extracted features

Knowing that OSM data are of low quality (e.g. green areas) or might not be up-to-date (e.g. buildings) we applied the semantic segmentation based on satellite images to be able identify these features more appropriately. Then they replaced OSM based features used in previous models. Finally, it concerns only buildings and green areas were identified with a different approach.

To gather features from satellite images, different neural network architectures were used. The results were compared with the usage of IoU metric (Intersection over Union) and based on that the final model was selected. The four different models were used (VGG16, VGG19, ResNet34, ResNet101) as the backbones in semantic segmentation and it appeared that, based on the IoU metric, the ResNet101 gives the best results. Sample output of the model mapped on the image is presented in Figure 5. It is visible that the algorithm is capable of segmenting buildings. It works the best for oblong buildings. In case of green areas, the selection of appropriate ranges of color channels is performed. Such technique was used as these areas are not marked in OSM data. In order to be able to create CNN solution for green areas, manual labelling needs to be performed, which is too time-consuming.

**Figure 5: Ground truth and semantic segmentation output**



Extracted features (area of buildings, area of greens) replaced the ones based on OSM. Comparison of these features is presented in Table 5. For predicted area of buildings, the coefficient variation is bigger than in OSM and is equal 61%. In total, solution based on satellite imagery gives 17% more of buildings area than OSM. In case of green areas, the CV decreased from 125% to 81% and in total the predicted values are 5 times greater. The remaining features in the model were left unchanged. It is because, the labels for buildings were collected from OSM and there is a lot of them so it is possible to build a solution recognizing buildings on images based on that. In case of roads, malls or public transportation stops, bigger training set will be required. In addition, semantic segmentation is a complex approach and creating solution for recognition of multiple objects requires a lot more of computational power.

**Table 5. Comparison of OSM features and features based on satellite images**

District	OSM_area_buildings	Predicted_buildings	OSM_area_green	Predicted_green
<b>PRAGA-POLUDNIE</b>	3008345	5343695	902701	9250267
<b>BIAŁOLEKA</b>	4355817	1242497	8606595	40295641
<b>MOKOTÓW</b>	4265839	2967480	3548119	15602519
<b>BIELANY</b>	2225264	5020159	1974975	12848376
<b>OCHOTA</b>	1774250	649568,7	502269	3779491
<b>BEMOWO</b>	2070012	1893136	845219	10506628
<b>REMBERTÓW</b>	1123002	4352027	408214	9781372
<b>TARGÓWEK</b>	2599218	5564022	1391163	10142156
<b>WAWER</b>	4201548	673209,8	7087154	37291578
<b>URSYNÓW</b>	3385563	1813806	3783516	21019126
<b>ŚRÓDMIEŚCIE</b>	3093402	3135141	430033	5445245
<b>WESOŁA</b>	1123001	2724055	24723	9793659
<b>URSUS</b>	1274779	1917677	142577	4358833
<b>PRAGA-PÓLNOC</b>	1368982	804254,6	628274	4468920
<b>WOLA</b>	2687147	4311564	781831	7133591
<b>WILANÓW</b>	1652998	3087434	8188930	16546570
<b>WŁOCHY</b>	2767635	5283471	975640	11188572
<b>ŻOLIBORZ</b>	974107	829339,3	474750	3699103
<b>mean</b>	2441717	2867363	2260926	12952869
<b>CV</b>	45%	61%	125%	81%

Again, two models, i.e. LASSO and random forest were used to estimate 9 models for each of the economic indicators as a dependent variable. The results of LASSO estimation in terms of variable importance ranking is presented in Table 6. The comparison with the model with OSM features only shows that satellite-based features have similar importance to OSM based ones. In general, 3 models (for population total, population men, life expectancy) achieve slightly better MAPE, but the difference in all 9 models is marginal.

Life expectancy model achieved the same MAPE as model based only on OSM data. Once again, public transportation and variables related to malls and supermarkets were the most important. Area of buildings and greens were at fifteenth and sixteenth place, which is similar to the previous model. In case of the model for income of share in PIT/CIT, the MAPE is higher for over 40% in comparison to OSM based only model. The predicted area of buildings was one of the least important features, which can indicate that this variable is noisy.

In the case of Random Forest models, predicted area of greens was the second most important feature in life expectancy model and sixth in income total model. Unfortunately, the predicted area of buildings was one of least important features. Generally, 3 (income total, vehicles total, vehicles passenger) models get improved in comparison with previous random forest models. The biggest improvement has been on income total with MAPE dropped from 135% to 126%. In comparison to the LASSO model these results look more promising as the predicted area of green is more important and the overall performance is better.

The research hypotheses were verified by performing multiple models with different dataset specifications. To get answer whether it is possible to predict socio-economic indicators, models based on OSM were performed (see Tables 3 and 4). In case of checking if features extracted from satellite images can replaces OSM data, models with such data were evaluated (see Tables 6 and 7). The last hypotheses were verified by comparison of LASSO and random forests.

Table 6. Ranking of feature importance and error measures for LASSO models based on OSM and extracted features

variable	budget2019_mln	income_share_in_PIT_CIT	income_total	population_total	population_women	population_men	vehicles_total	vehicles_passenger	life_expectancy
predicted_area_buildings		15							15
predicted_area_green							9	9	16
OSM_area_water		13	10						14
OSM_number_bikeshare_stations		2	2	4	4		1	2	6
OSM_number_buildings									13
OSM_number_fuel_stations		6	4	3	1	1	4	4	8
OSM_publictrans_stops_bus				6	3	2			7
OSM_publictrans_stops_subway		3	3	2					3
OSM_publictrans_stops_train		4					3	3	2
OSM_publictrans_stops_tram		7		5	2		5	5	4
OSM_shops_number_mall		1					2	1	5
OSM_shops_number_supermarket		5	1	1					1
OSM_streets_length_primary		12	5				8	8	11
OSM_streets_length_residential		11							
OSM_streets_length_secondary		10	8	7	5	3	6	6	9
OSM_streets_length_tertiary		9	7						
OSM_streets_length_trunk		8	6				7	7	12
OSM_streets_length_unclassified		14	9						10
MAPE	64,9%	102,6%	84,6%	52,6%	63,2%	63,3%	57,0%	59,8%	5,6%

Table 7. Ranking of feature importance and error measures for random forest models based on OSM and extracted features

variable	budget2019_mln	income_share_in_PIT_CIT	income_total	population_total	population_women	population_men	vehicles_total	vehicles_passenger	life_expectancy
predicted_area_buildings	13	16	16	14	14	14	12	14	16
predicted_area_green	15	14	6	16	16	16	16	16	2
OSM_area_water	8	12	11	12	12	12	10	9	4
OSM_number_bikeshare_stations	1	2	2	1	1	1	2	2	1
OSM_number_buildings	11	10	12	6	7	6	9	10	6
OSM_number_fuel_stations	5	7	10	3	3	4	4	3	12
OSM_publictrans_stops_bus	4	9	7	5	5	5	7	7	13
OSM_publictrans_stops_subway	12	4	4	15	15	15	11	11	15
OSM_publictrans_stops_train	17	11	14	17	17	17	14	13	3
OSM_publictrans_stops_tram	7	3	3	11	9	11	5	4	5
OSM_shops_number_mall	3	5	9	4	4	3	8	8	17
OSM_shops_number_supermarket	14	17	13	9	10	9	17	17	11
OSM_streets_length_primary	6	1	1	8	8	8	3	5	9
OSM_streets_length_residential	16	13	17	13	13	13	13	12	7
OSM_streets_length_secondary	9	6	5	7	6	7	1	1	10
OSM_streets_length_tertiary	2	8	8	2	2	2	6	6	14
OSM_streets_length_trunk	18	18	18	18	18	18	18	18	18
OSM_streets_length_unclassified	10	15	15	10	11	10	15	15	8
MAPE	40,0%	65,1%	126,7%	52,4%	53,7%	51,0%	66,3%	69,4%	2,2%

## 6. Conclusion

The main goal of the project was to verify whether features visible from space can be used to predict socio-economic indicators on the local level. Three research hypotheses have been tested. The first one, concerning possibility of predicting various socio-economic indicators with the features visible from space has been verified by creating models based on the OSM data. In the second hypothesis, quality of features extracted from images were checked. It was done by comparing models based on different datasets. The last hypothesis assumes that machine learning models can achieve better accuracy in prediction of well-being indicators than linear models. It was checked by performing comparison of random forest and LASSO.

Analyzing all the results, it is visible that best accuracy is achieved for population, vehicles and life expectancy. In the case of life expectancy such a result is not surprising as it does not differ very much among districts. The lowest error was achieved for random forest solution based on OSM and extracted features from satellite images. For income share in PIT/CIT the best model was LASSO based on OSM data, where only three variables were taken into consideration. In general, the predictive capability of OSM and satellite imagery features is confirmed on a local level, which confirms the first hypothesis of the research.

One of the challenges in the process is the fact that semantic segmentation and in general computer vision tasks require large computing power and good training data. In approach presented in the article, the labels of objects are needed, so the OSM data was used. The limitations and inaccuracies of such solution were presented above, but manual labeling was too time-consuming to be taken into consideration.

In this research only two variables: the area of buildings and green areas are calculated based on satellite images and performance of models using these features is similar to models that base on OSM features only. Therefore, the second research hypothesis was not confirmed, but on the other hand its verification was limited. In order to obtain a wider view of satellite imagery usefulness, more features need to be extracted from satellite images, which in turn requires much more computational power and huge effort of preparing well labeled training dataset.

In total 36 (18 specifications for LASSO and random forest) models were estimated. It comes out that a random forest, which is more robust and capable of catching non-linearities, achieved better results in 11 out of 18 specifications. Such results confirm the third research hypothesis that machine-learning algorithms offer better accuracy than linear models.

There are some limitations and possible further steps that should be highlighted to improve the current solution. In order to get better local measurements, well-being data should be on a level lower than a district, especially if the analysis is based on one city which is rather heterogenous and specific in Poland. Unfortunately, the data on which the research is based is on the lowest level of aggregation that is publicly available. Furthermore, the objective of the model is a regression task (predicting a continuous variable), which is relatively more difficult than classification (predicting a categorical variable with a limited number of levels).

Another idea of future extension would be to analyze data over time to observe changes. It would enable prediction of socio-economic indicators for the upcoming years which could be helpful for policy makers. In addition, such models could be used on areas that have worse labelled data. High-resolution satellite images with even daily frequency are already available, however they are still relatively costly.

To sum up, research hypotheses were verified and it comes out that a solution based on OSM data in combination with satellite images and machine learning tools is capable of predicting socio-economic indicators on local level with decent quality. There are some limitations regarding the quality of the data and challenging process of training neural networks, but our results show potential for performing such research for the area of Poland at the local level.

## References

- Babenko, B., Hersh J., Newhouse D., Ramakrishnan A., & Swartz T. (2017), "Poverty Mapping Using Convolutional Neural Networks Trained on High and Medium Resolution Satellite Images, With an Application in Mexico", 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA. <https://arxiv.org/abs/1711.06323>
- Bennett, M. M., & Smith, L. C. (2017), "Remote Sensing of Environment Advances in using multitemporal night-time lights satellite imagery to detect , estimate , and monitor socioeconomic dynamics", *Remote Sensing of Environment*, vol. 192, pp. 176–197. <https://doi.org/10.1016/j.rse.2017.01.005>
- Berrar, D. (2018), "Cross-validation", Data Science Laboratory, Tokyo Institute of Technology.
- Bickenbach, F. Bode, E., Nunnenkamp P., & Soder M. (2016), Night lights and regional GDP. *Review of World Economics*, vol. 152(2), pp. 425–447.
- Chen, Y., Liu, X., Li, X., Liu, Y. & Xu, X. (2016), "Mapping the fine-scale spatial pattern of housing rent in the metropolitan area by using online rental listings and ensemble learning.", *Applied Geography* 75 (2016) 200-212. <https://doi.org/10.1016/j.apgeog.2016.08.011>
- Chollet, F. (2018), "Deep Learning with Python", Manning Publications co.
- Ciołek, D. (2017), „Oszacowanie wartości produktu krajowego brutto w polskich powiatach”, *Gospodarka Narodowa*, nr 27 (3).
- Elvidge, Ch. D., Sutton P. C., Ghosh, T., Tuttle, B. T., Baugh, K. E., Bhaduri, B. & Bright E. (2009), "A global poverty map derived from satellite data", *Computers & Geosciences*, Volume 35, Issue 8, pp. 1652-1660. <https://doi.org/10.1016/j.cageo.2009.01.009>
- Elvidge, C.D., Baugh, K., Anderson, S. & Suttion, P. C., (2012), "The Night Light Development Index (NLDI): a spatially explicit measure of human development from satellite data", *Social Geography*, vol. 7(1), pp. 23–35. doi:10.5194/sg-7-23-2012
- Engstrom, R, Hersh, J. & Newhouse, D. (2017), "Poverty from space: using high-resolution satellite imagery for estimating economic well-being", Policy Research working paper; no. WPS 8284; Paper is funded by the Strategic Research Program (SRP). Washington, D.C.: World Bank Group.

- <http://documents.worldbank.org/curated/en/610771513691888412/Poverty-from-space-using-high-resolution-satellite-imagery-for-estimating-economic-well-being>
- Gennaioli, N., La Porta, R., De Silanes, F. L. & Shleifer A. (2014), “Growth in Regions”, *Journal of Economic Growth*, vol. 19(3), pp. 259-309.
- He K, Zhang, X., Ren, S. & Sun, J. (2015), “Deep Residual Learning for Image Recognition”. <https://arxiv.org/abs/1512.03385>
- Henderson, J.V., Storeygard, A. & Weil, D. N. (2012), “Measuring Economic Growth From Outer Space”, *The American economic review*, vol. 102(2), pp. 994–1028. DOI: 10.1257/aer.102.2.994
- Geron, A. (2018, p.141-143, 351-353), “Uczenie maszynowe z użyciem Scikit-Learn i Tensorflow”, Helion S.A.
- Goodfellow, I. et al. (2016), “Deep Learning”, MIT Press.
- Ioffe S. and Szegedy C., (2015), “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift”. <https://arxiv.org/abs/1502.03167>
- Jasiński, T. (2019), “Modeling electricity consumption using nighttime light images and artificial neural networks”, *Energy*, vol. 179, pp. 831–842, doi: <https://doi.org/10.1016/j.energy.2019.04.221>
- Jean, N., Burke, M., Xie, M., Davis, W.M., Lobell, D.B. & Ermon, S. (2016), “Combining satellite imagery and machine learning to predict poverty”, *Science*, vol. 353, issue 6301, pp. 790-794. doi: 10.1126/science.aaf7894
- Jordan J. (2018), “An overview of semantic image segmentation” <https://www.jeremyjordan.me/semantic-segmentation/>
- Lessmann, Ch. and Seidel A. (2017), “Regional inequality, convergence, and its determinants – A view from outer space”, *European Economic Review*, vol. 92, issue C, pp. 110-132.
- Li F., Johnson J., and Yeung S., (2017), CS231n: Convolutional Neural Networks for Visual Recognition, Lecture 11: Detection and Segmentation, course materials on Stanford university, <http://cs231n.stanford.edu/>.
- Mellander, C., Lobo, J., Stolarick, K. & Matheson, Z. (2015), “Night-time light data: A good proxy measure for economic activity?”, *PLoS ONE*, vol. 10(10), pp. 1–18. <https://doi.org/10.1371/journal.pone.0139779>
- Nordhaus, W. D. and Chen, X. (2011), “Using luminosity data as a proxy for economic statistics”, *Proceedings of the National Academy of Sciences*, vol. 108(21), pp. 8589–8594.

- Pinkovskiy, M. and Sala-I-Martin, X. (2016), “Lights, Camera ... Income! Illuminating the National Accounts-Household Surveys Debate”, *The Quarterly Journal of Economics*, vol. 131(2), pp. 579–631.
- Pomianek, I. and Chrzanowska, M. (2016), „A spatial comparison of semi-urban and rural gminas in Poland in terms of their level of socio-economic development using Hellwig’s method”, *Bulletin of Geography. Socio-economic Series*, vol. 33 (33), pp. 103-117.
- Ronneberger, O., Fischer, P. & Brox, T. (2015), “U-Net: Convolutional Networks for Biomedical Image Segmentation”. <https://arxiv.org/abs/1505.04597>
- Simonyan K. and Zisserman A., (2015), “VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION”, Published as a conference paper at ICLR 2015.
- Sompolska-Rzechuła A. (2016), „Zróźnicowanie rozwoju społecznego w ujęciu przestrzennym”, *Wiadomości statystyczne*, no 1, pp. 62-78.
- Tibshirani R. (1996), “Regression shrinkage and selection via the Lasso”, *Journal of the Royal Statistical Society Series B*, vol. 58 (1), pp. 268-288.
- Tingzon, I., Order, A., Sy, S., Sekara, V., Weber, I., Fatehkia, M., Herranz, M. G. & Kim, D. (2019), “Mapping Poverty in the Philippines Using Machine Learning, Satellite Imagery, and Crowd-sourced Geospatial Information“, *International Conference on Machine Learning AI for Social Good Workshop*, Long Beach, United States, 2019.
- UNDP (2012), „UNDP National Human Development Report. Poland 2012. Local and Regional Development”.
- Wójcik P. (2018b), „Nighttime light as a proxy for economic development case of Poland”, V *International Scientific Conference Spatial Econometrics and Regional Economic Analysis – SEREA 2018*, University of Lodz, Łódź, 14–15 June.

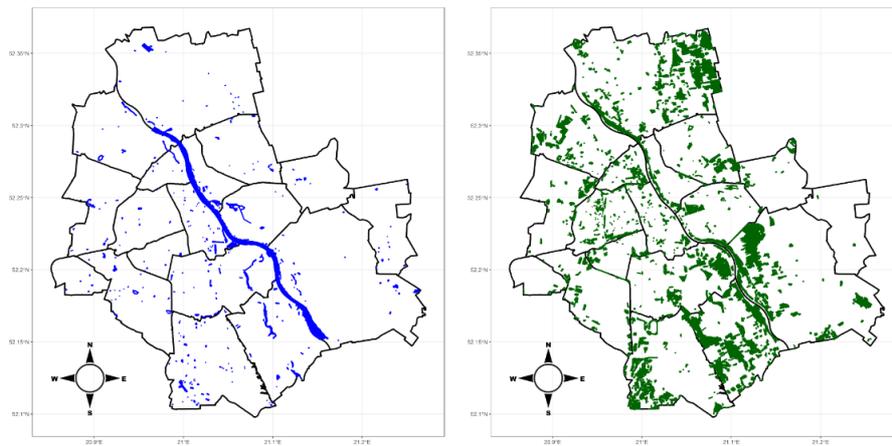
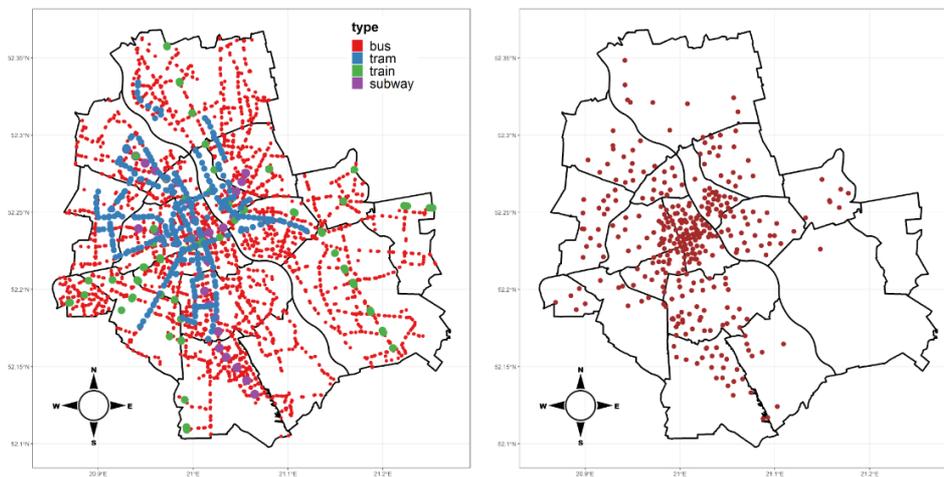
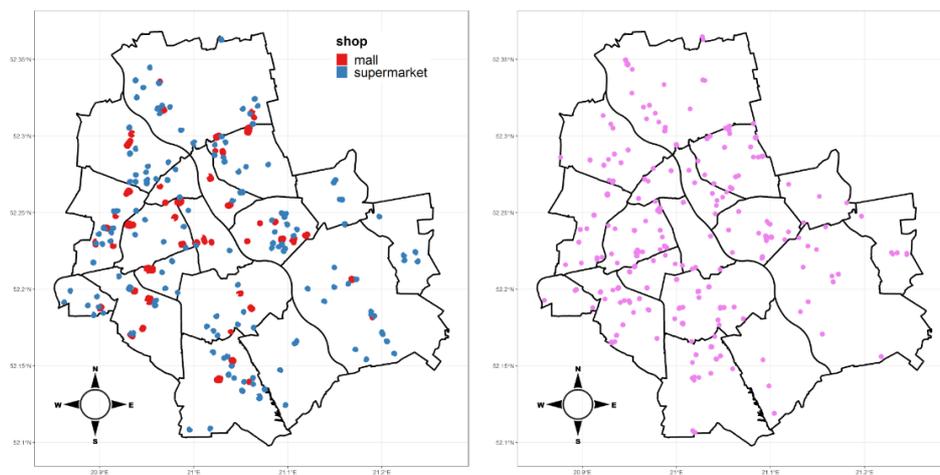
## Appendix

Figure A1. Open Street Map features mapped on part of Warsaw



Figure A2. OSM - buildings and roads



**Figure A3. OSM - water and green areas****Figure A4. OSM - public transportation stops and bikeshare points****Figure A5. OSM - shops (malls, supermarkets) and fuel stations**

**Figure A6. Sample OSM based labels of buildings**



**Figure A7. Alternative ways Extracting green areas from satellite images – original image (left), green channel extraction (middle), extraction based on OSM data (right)**



Figure A8. Scatter plot of income share in PIT/CIT and number of bikeshare stations

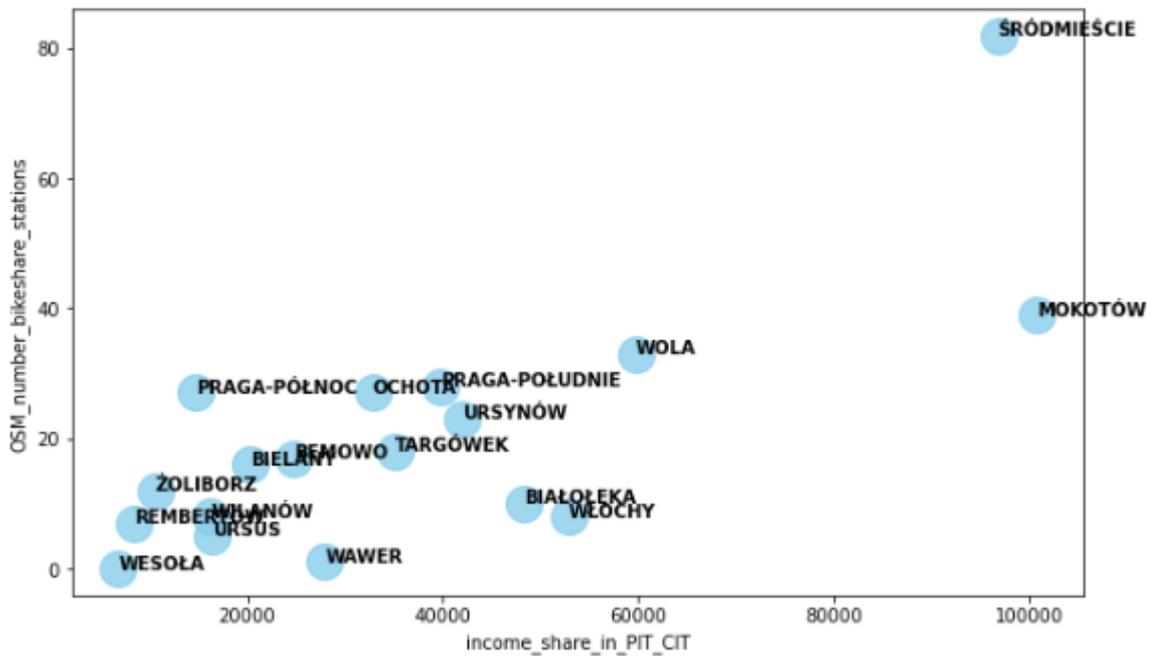


Figure A9. Scatter plot of income share in PIT/CIT and area of building

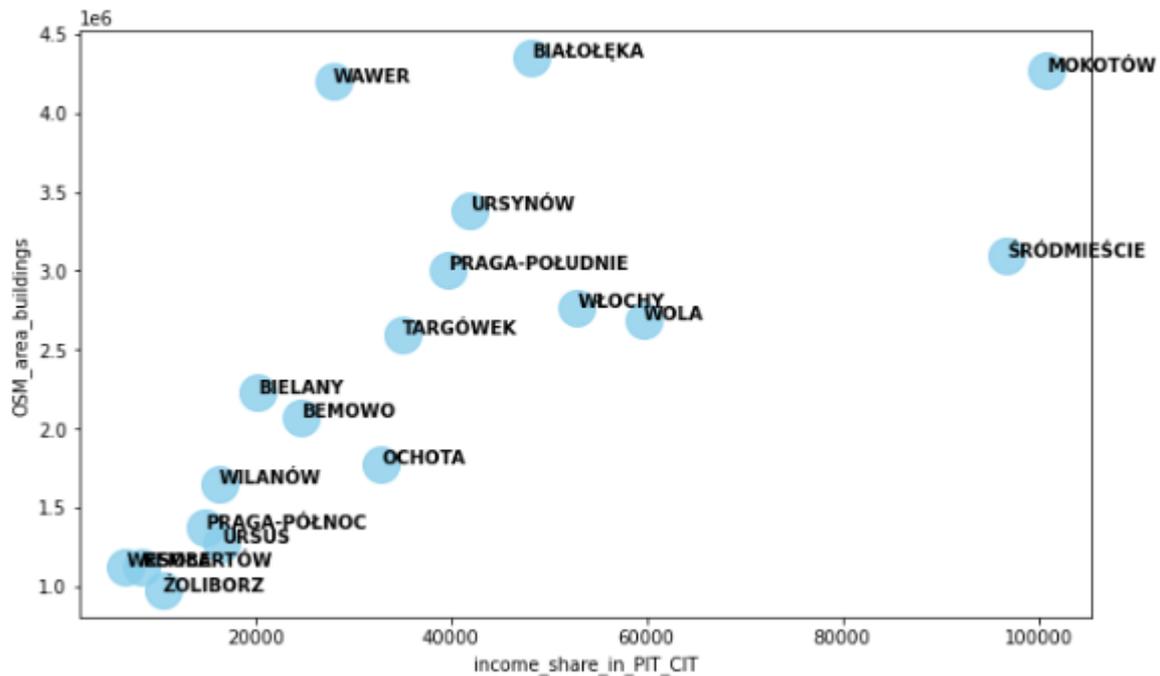


Figure A10. Scatter plot of income share in PIT/CIT and area of building

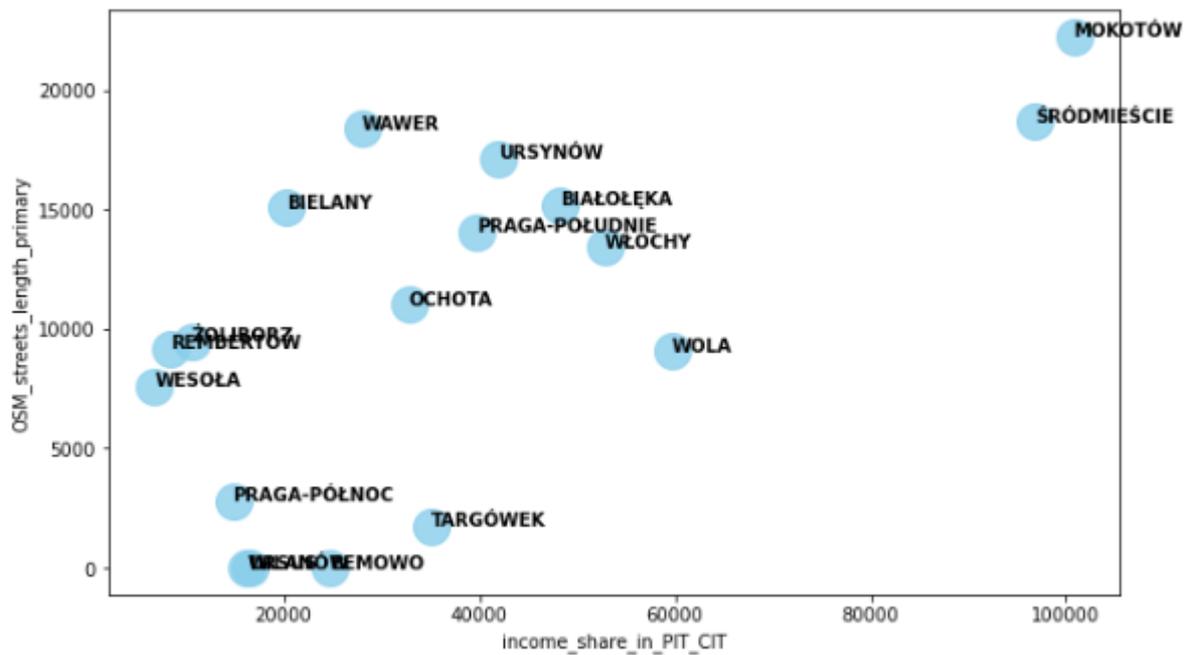


Figure A11. Scatter plot of life expectancy and number of bikeshare stations

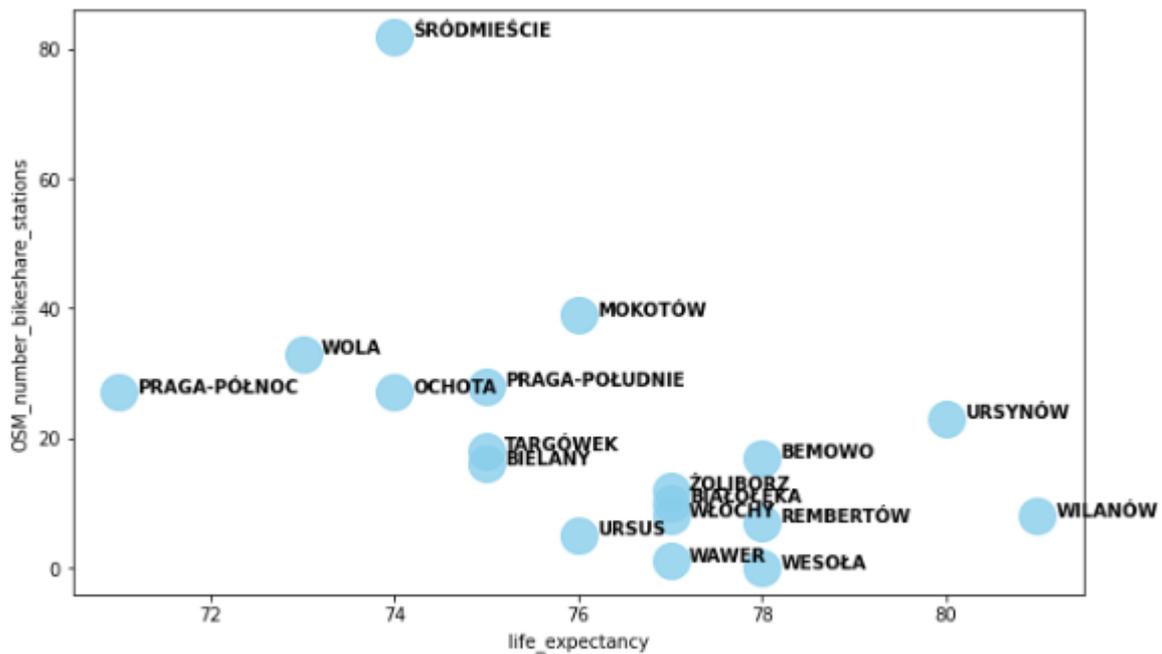


Figure A12. Scatter plot of life expectancy and area of green

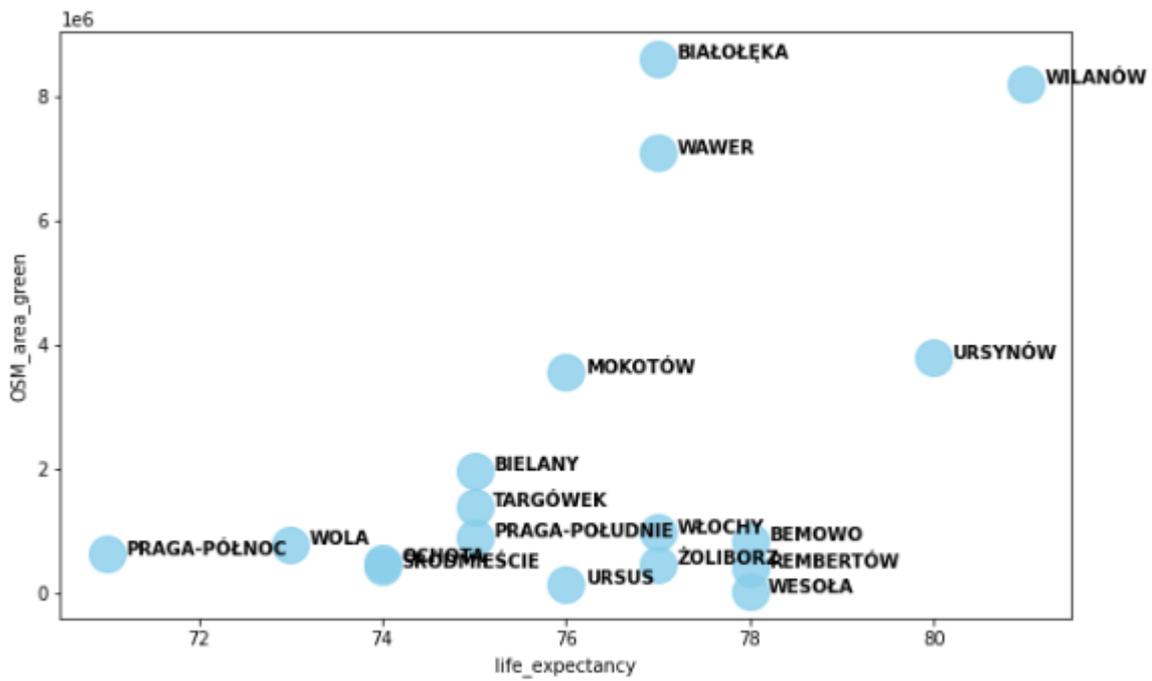
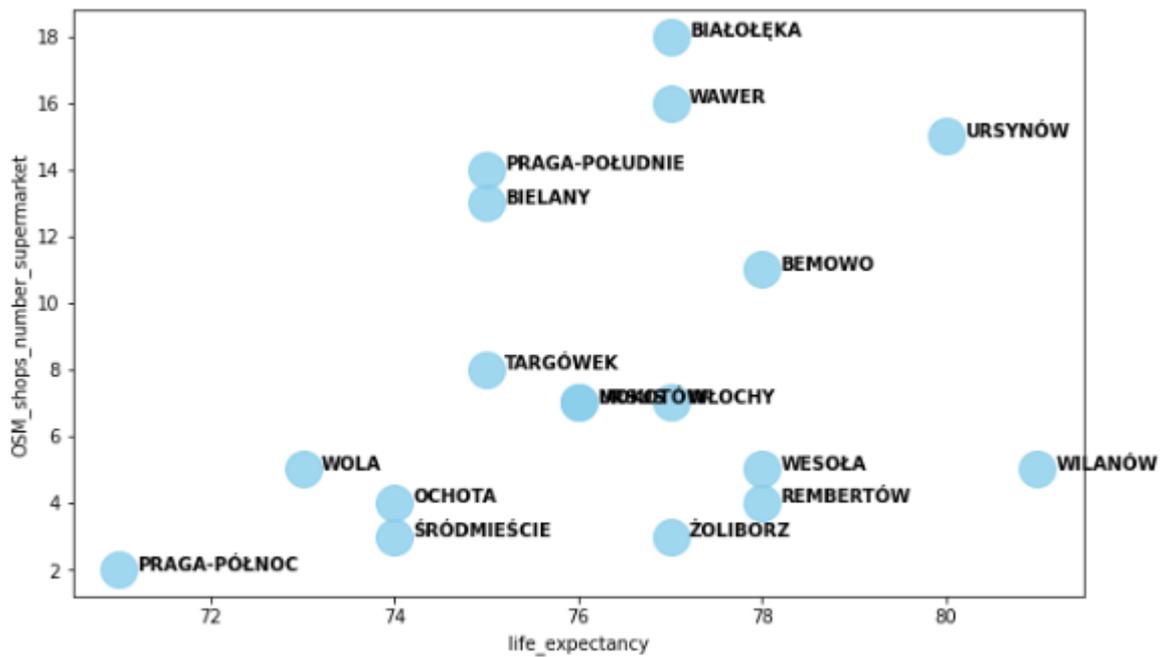


Figure A13. Scatter plot of life expectancy and number of supermarkets





UNIVERSITY OF WARSAW

FACULTY OF ECONOMIC SCIENCES

44/50 DŁUGA ST.

00-241 WARSAW

[WWW.WNE.UW.EDU.PL](http://WWW.WNE.UW.EDU.PL)