academic publishers
INTERNATIONAL JOURNAL OF ARTIFICIAL INTELLIGENCE (ISSN: 2692-5206)
Volume 04, Issue 04, 2024
Published Date: - 14-06-2024

INNOVATION IN COMPANY LABOR PRODUCTIVITY MANAGEMENT: DATA SCIENCE METHODS APPLICATION

Khaydarova M.Sh.
PhD student at TSUE100066, Uzbekistan
https://orcid.org/0009-0003-5287-7787
Email:mkhaydarova93@gmail.com

Abstract: The article addresses the challenge of enhancing labor productivity in companies by analyzing objective data on economic, demographic, and social factors alongside subjective information about employees' health quality. It introduces a technology for labor productivity management that involves phased data processing and modeling relationships between quantitative and qualitative data, aimed at aiding decision-making for planning productivity growth trajectories. This technology leverages statistical analysis and machine learning to support managerial decisions on health-preserving strategies to boost productivity. It is demonstrated that the k-means method is more suitable and reliable for clustering employees into homogeneous groups compared to Kohonen neural networks. Additionally, various classification methods for predicting new employees' labor productivity profiles were tested, revealing that support vector machines offer higher accuracy when working with numerous qualitative variables such as gender, education, and health self-assessment.
Keywords: data science; statistical data processing; predictive analytics; machine learning; classification; clustering; labor productivity; health management; health-saving strategies; electric power industry

1. Introduction

Rapid modernization and technological innovations underscore the critical issue of human capital development, with the quality of human capital, including health, significantly impacting labor productivity growth. The study's core idea is to emphasize the importance of investing in personnel health to enhance a company's labor productivity, as well as to develop new technologies and models for implementing such decisions. Improving labor productivity and modernizing key economic sectors are essential for economic development and national security. Labor productivity is a vital indicator of a country's economic progress and a determinant of the standard of living. The Russian national project "Labor productivity increasing and employment support" aims to enhance productivity in companies. Addressing issues such as improving labor productivity, boosting companies' performance and competitiveness, and fostering industry and regional development relies heavily on maintaining and ensuring human health and safety.

Labor productivity, an economic measure of labor and production efficiency, indicates the value of labor costs needed to produce a unit of product.Traditionally, the factors that contribute to labor productivity improvement [2–5] are grouped as follows: - Material and technical factors, including technological innovations; - Establishment of high-performance workspaces; - Development, specialization, and concentration of production, along with the implementation of lean production techniques; - Enhancement of production structure and output volumes; - Advanced training; - Social and economic factors, such as determined wages and working conditions.

As indicated above, human potential factors such as employee health quality are not typically considered productive or targeted [6–8]. For a thorough analysis of this issue, there is a lack of appropriate methods for quantitatively assessing the impact of personnel health on labor productivity, which is essential for decision support and strategy creation in health management to increase labor productivity and enhance a company's operational efficiency. We define human health according to the World Health Organization (WHO) [9] as "a state of complete physical, mental, and social well-being, and not just the absence of diseases and physical defects." Health status, social skills, and knowledge reflect human potential. Improving the quality of human potential directly influences the performance metrics of companies, organizations, and economies overall. The state and quality of human health as a factor in workforce productivity have not been sufficiently explored in scientific research. This issue will likely become more prominent once other mechanisms and sources of labor efficiency growth, such as production modernization, digitalization, and organizational improvements, have been exhausted. Amid the fourth industrial revolution, workforce quality is shifting from being a minor to a crucial factor in labor productivity management.

The issue of personnel health protection in human resource management is discussed in [10–12], where the current labor protection quality management system is analyzed and mechanisms for improvement are proposed. Health factors are also considered when developing quality management systems [13]. Health, viewed as part of human capital and health preservation as an element of corporate social responsibility, is explored in [14–17]. The effectiveness of healthcare systems at the macroeconomic level, the impact of health preservation investments on economic growth, and their overall effectiveness are discussed in [18–29]. Qualitative analyses based on expert assessments of the impact of personnel health on labor productivity are conducted in [30,31] and are regularly featured in WHO and OECD reports [8,9,32]. Numerous studies focus on approaches to labor productivity growth as a factor of company competitiveness, including methods for talent search, training, and promotion [33–36]. Leading indicators that characterize labor conditions and the value-motivational environment are presented in [37–39]. The economic returns to health and its impact on employment and wages are discussed in [40,41]. Studies on the influence of employee age on labor demand, considering the decrease in health potential with age, are presented in [42]. Several publications view enterprises as complex social systems, exploring the interdependent relationship between organizational effectiveness, employee health, and quality culture [43]. According to the systematic view of work organizations, employee health is closely tied to organizational effectiveness, and building a healthy organizational culture is crucial for promoting both [44–46]. Considering employers' interests in improving employee health to boost company efficiency, several studies examine the impact of health measures on the future labor market outcomes of employees [47].

Through analyzing existing approaches, methods, and models of personnel health and labor productivity management, several significant shortcomings have been identified, limiting their applicability. Notably, there are no methods for quantitatively assessing how health levels and conditions impact labor productivity, nor are there recommendations for forming comprehensive management decisions aimed at improving labor resource efficiency while considering the quality of these resources. The challenge lies in studying the diverse factors affecting labor productivity, including employee health, and developing models to reveal the nature and type of relationships between these factors and identifying homogeneous employee groups with similar productivity profiles for better management. This is particularly relevant and meaningful amid the digital transformation of the economy. Hence, there is a need for a new approach, technology, and supporting models that reflect the essential properties of enterprises' socio-economic systems, characterized by the high dynamics of ongoing processes and the uncertainties of internal and external environments. This work hypothesizes that for a more complete and comprehensive description of labor productivity as a management object, beyond economic factors, it is essential to consider social, demographic, and health-related factors. The study aims to justify the feasibility of financial investment in maintaining employee health to ensure labor productivity growth in enterprises, as well as to design the technology and models supporting such investment.

The objective of this study is to develop a technology for labor productivity management in a company by considering heterogeneous information related to economic, demographic, and social factors, as well as the quality factors of personnel health, to support decision-making in planning labor productivity growth trajectories. To achieve this objective, the following problems are addressed: - Identifying and substantiating a set of factors, including health quality factors, that determine labor productivity. -

Designing homogeneous employee groups using the aforementioned factors. - Developing a set of management decisions tailored to improve personnel health quality for each homogeneous group, thereby contributing to labor productivity growth. - Assessing the economic efficiency of the proposed management decisions aimed at preserving personnel health. It is important to note that this study does not cover issues related to occupational health and safety or the reduction of occupational injury risks. Instead, it focuses on factors that reflect employees' health potential and their capability to perform professional duties, as well as the employers' intent to influence this potential to maintain, enhance, and support labor productivity growth. The methodology of the research includes human capital theories, methods of system analysis and modeling of social and economic processes, statistical modeling methods, cluster analysis methods, and decision-making methods under complexity. The research materials consist of statistical data and operational reports from a large company in the electric power industry, as well as data from employee surveys conducted within the company.

2. Methodology of Research

The proposed technology for labor productivity management involves step-by-step data processing and modeling, taking into account demographic, social factors, and personnel health quality. Below is the conceptual diagram and implementation stages: **Stage 1: Qualitative Analysis of Personnel Health Status, Technical and Economic Data Acquisition** - Conduct continuous examinations of the company's employees through medical exams and questionnaires. - Gather objective data on employees' health status and subjective information regarding their health state and quality. **Stage 2: Selection and Substantiation of Factors Affecting Individual Productivity** - Perform an exploratory data analysis to assess the impact of various factors. - Use correlation and regression analysis to evaluate the influence of social and demographic characteristics, as well as health state and quality, on labor productivity. **Stage 3: Employee Clustering into Homogeneous Groups** - Group employees with similar demographic, social factors, and health quality into homogeneous clusters. - Develop typical personnel profiles for each cluster. **Stage 4: Development of Management Decisions** - Create targeted management decisions aimed at improving the health state and quality for each homogeneous employee cluster. The cumulative aim of these stages is to provide a comprehensive approach to labor productivity management by factoring in various influences, especially those related to health, to ensure optimal decision-making and productivity growth. Is there anything specific you would like to delve deeper into or any other assistance you need related to this technology development?

Verification of Hypothesis and Technology

The hypothesis and developed technology were tested using empirical data from a large electric power company in Russia. An experiment conducted in 2020 involved over 700 employees from the technical, planning, and financial departments. Data collection was performed through a comprehensive survey, detailed in Appendix A (Table A1). A qualitative analysis of employees' health status was conducted through a developed questionnaire consisting of 30 questions, enabling self-assessment of their health. The results were processed based on five indicators characterizing different health conditions, as shown in Appendix A (Table A2). The survey results were evaluated by four experts.

Data Collection and Assessment

In addition to the survey data, objective information such as employees' education, marital status, and number of children was gathered from the company's databases. The calculation of labor productivity followed the algorithm described in Section 3.1. Data analysis and modeling were performed using Statistica 10.0 software.

Machine Learning Methods

Statistical machine learning methods, part of data science, differ from classical statistical methods in that they are data-driven and do not necessarily describe the data using a linear or other general function.

Machine learning focuses on developing efficient algorithms that scale to large datasets to optimize predictive models. Below is a brief description of the machine learning methods employed in this study:

1. **Exploratory Data Analysis (EDA)**: Used to understand data distributions and relationships, serving as a preliminary step to apply more advanced techniques.
2. **Correlation and Regression Analysis**: Helps identify and quantify the impact of various factors on labor productivity.
3. **Clustering (e.g., K-means)**: Groups employees with similar demographic, social, and health characteristics into homogeneous clusters.
4. **Support Vector Machines (SVM)**: A classification method used to predict labor productivity profiles based on various qualitative variables.

2.2.1. Decision Tree Models in Data Science

Decision tree models are robust classification tools in data science, known for their predictive accuracy and interpretability. The core of a decision tree model lies in recursive partitioning, which divides data into sections and subsections to create homogeneous classes in each subset. This process results in a set of "if-then-else" implication rules, capable of uncovering hidden patterns within complex data interactions.
Intuitive Algorithm
The recursive partitioning algorithm for building a decision tree is intuitive. Data is divided multiple times based on predictor values, resulting in relatively homogeneous segments. This division continues until a predetermined stopping criterion is met.
Types of Decision Trees
Various top-down decision tree inducers include:
ID3: Focuses on information gain for splitting data.
C4.5: An extension of ID3, incorporates both growing and pruning phases to avoid overfitting.
CART (Classification and Regression Trees): Uses Gini impurity or entropy for classification trees and variance reduction for regression trees. It also includes growing and pruning phases.
These algorithms are detailed in resources such as [51].

2.2.2. Support Vector Machines (SVM)

Support Vector Machines (SVM) are powerful supervised learning methods used for both classification and regression analysis. The core idea behind SVM is to construct a hyperplane that optimally separates the sample objects into distinct classes. The greater the distance or margin between the hyperplane and the sample objects of different classes, the lower the average classification error.

Advantages of SVM:

- **Effective in High-Dimensional Spaces:** SVMs are particularly effective when dealing with large feature spaces.
- **Efficient with Large Dimensions:** They work well even when the number of features exceeds the number of samples.
- **Memory Efficiency:** SVM uses support vectors (a subset of the training set) in the decision function, making it memory efficient.
- **Versatility:** SVM allows for different kernel functions (linear, polynomial, radial basis function, sigmoid, etc.) to be used in the decision function. Custom kernels can also be specified.

Disadvantages of SVM:

- **Risk of Overfitting:** When the number of features is much larger than the number of samples, selecting the right kernel features and applying regularization is crucial to avoid overfitting.

- **No Direct Probability Estimates:** SVMs do not inherently provide probability estimates; calculating probabilities requires an additional step, often using expensive five-fold cross-validation.

2.2.3 K-means Clustering Method

K-means is a widely used clustering method that partitions a set of data points in a vector space into **k** predetermined clusters. Here's a detailed breakdown of how the algorithm works:

Initialization: Select **k** initial centroids randomly from the data points.

Assignment Step: Assign each data point to the nearest centroid based on the chosen metric (e.g., Euclidean distance).

Update Step: Recalculate the centroids by computing the mean of all data points assigned to each cluster.

Iteration: Repeat the assignment and update steps until convergence is reached. Convergence occurs when:

There is no change in the intra-cluster distance, meaning data points remain in the same clusters as in the previous iteration.

A finite number of iterations is reached, ensuring the algorithm won't loop indefinitely due to the finite number of partitions and decreasing total standard deviation.

The algorithm aims to minimize the sum of squared distances between each data point and the centroid of the cluster to which it is assigned. This results in a set of compact, well-defined clusters.

Key Characteristics:

- The algorithm iteratively recalculates the center of mass (centroid) for each cluster.
- It reassigns data points to the clusters based on the updated centroids.
- The process continues until a stable cluster configuration is achieved.

Advantages of K-means:

- Simple and easy to implement.
- Efficient for large datasets.
- Adaptable to new examples (centroids can be updated with incremental changes).

Disadvantages of K-means:

- Requires the number of clusters **k** to be specified in advance.
- Sensitive to the initial placement of centroids.
- Prone to converging to local minima.
- Not suitable for clusters with irregular shapes or varying densities.

2.2.4 Self-Organizing Maps (SOM)

Self-Organizing Maps (SOM) are a type of unsupervised neural network algorithm used for clustering and dimensionality reduction. Unlike neural networks trained with the backpropagation algorithm, SOMs use unsupervised learning, meaning the training result depends solely on the structure of the input data.

Key Characteristics:

- **Unsupervised Learning:** The algorithm learns without labeled output data, making it suitable for exploring the underlying structure of data.
- **Multi-Dimensional Clustering:** SOM is an effective method for clustering high-dimensional data.
- **Topology Preservation:** SOM neurons (or nodes) are ordered into a structured grid, often two-dimensional. During training, not only is the winning neuron (the neuron whose weights are closest to the input vector) adjusted, but also its neighboring neurons to a lesser extent.

Functioning of SOM:

1. **Initialization:**
   - Initialize the weight vectors of the neurons, usually with small random values.
2. **Training:**
   - For each input vector:
     1. **Find the Best-Matching Unit (BMU):**
        - Identify the neuron whose weight vector is closest to the input vector.
     2. **Update Weights:**
        - Adjust the weights of the BMU and its neighboring neurons to be closer to the input vector.
        - The degree of adjustment decreases with distance from the BMU and over time (both the learning rate and neighborhood radius decrease).
3. **Iteration:**
   - Repeat for a set number of iterations or until convergence, where minimal changes occur in the map.

Advantages of SOM:

- **Dimensionality Reduction:** Projects high-dimensional data into a lower-dimensional space (e.g., 2D grid), preserving topological relationships.
- **Visualization:** The resulting maps provide intuitive and useful visual representations of complex data.
- **Clustering:** Naturally forms clusters based on data similarity, useful for exploratory data analysis.

Applications of SOM:

- Pattern recognition
- Data visualization
- Feature extraction
- Data compression

Comparison with K-means:

- While both SOM and K-means are clustering algorithms, SOM also organizes neurons into a grid and adjusts neighboring neurons, preserving the topological properties of the input space.

3. Experimental Results

To effectively utilize qualitative and quantitative indicators for our study, we need to organize and process these data systematically. Here's an outline of how we can manage and analyze these indicators:

Data Collection and Organization

Indicators to Collect:

1. **Social Characteristics:**
    - Education level
    - Marital status
    - Number of children
2. **Anthropometric Characteristics:**
    - Gender
    - Age
3. **Health Self-Assessment:**
    - Current health problems
    - Chronic diseases
    - Health self-assessment (self-reported health status)
    - Proper nutrition (diet quality)
    - Bad habits (e.g., smoking, alcohol consumption)

Data Analysis

1. **Descriptive Statistics:**
    - Calculate descriptive statistics (mean, median, mode, standard deviation) for each quantitative variable (age, number of children).
    - Summarize categorical data (education level, marital status) by frequency counts and percentages.
2. **Exploratory Data Analysis (EDA):**
    - Use correlation analysis to explore relationships between variables (e.g., the relationship between health self-assessment and age).
    - Use visualizations (e.g., histograms, bar charts, box plots) to understand the distribution and relationships of variables.
3. **Clustering Analysis:**
    - Apply K-means clustering or SOM to group employees into clusters based on their social characteristics, anthropometric data, and health self-assessment.
    - Evaluate the clustering results by examining cluster centroids and the distribution of data within clusters.

Implementation Steps

1. **Data Processing:**
    - Clean the data (handle missing values, normalize/scale features if necessary).
    - Encode categorical variables (e.g., marital status, education level) using one-hot encoding or label encoding.
2. **Model Building:**

    Use K-means:

```
from sklearn.cluster import KMeans
```

```python
kmeans = KMeans(n_clusters=k)
kmeans.fit(data)
labels = kmeans.labels_
```

Python

Use SOM with a library such as MiniSom:

```python
from minisom import MiniSom


som = MiniSom(x, y, input_len)
som.train_random(data, num_iterations)
```

3. **Analysis and Interpretation:**
   - Analyze the clusters to identify common characteristics within each group.
   - Develop targeted management decisions to improve the state and quality of personnel health based on the profiles of each cluster.

To effectively transition qualitative values into quantitative ones using binary coding and subsequently calculate individual labor productivity, we should follow a structured process. Here's how you can achieve that:

1. **Transition of Qualitative Values to Quantitative Values Using Binary Coding:**

- Each qualitative feature is converted to a binary representation (0 and 1), where 1 represents the presence or higher intensity of a characteristic, and 0 represents its absence or lower intensity.

Example Encoding:

- **Health Self-Assessment:**
  - Excellent: 1
  - Good: 0.75
  - Fair: 0.5
  - Poor: 0.25
  - Very Poor: 0
- **Proper Nutrition:**
  - Yes: 1
  - No: 0
- **Bad Habits:**
  - None: 0
  - Smoking: 1 (if present)
  - Alcohol: 1 (if present)

2. **Calculate Individual Labor Productivity:**

Individual labor productivity is determined using the value added per employee, a common measure in non-resource sectors of the economy.

Calculation Formula:

[\text{Individual LP} = \frac{\text{Value Added}}{\text{Number of Employees}}]

Collect relevant financial data from company records.

3. **Employee Survey:**

Collect data using developed questionnaires that gather all required qualitative and quantitative information.

Example Survey Questions:

- **Education Level:**
  - What is the highest level of education you have completed? [List educational levels]
- **Marital Status:**
  - What is your current marital status? [Single/Married/Divorced/Widowed]
- **Number of Children:**
  - How many children do you have?
- **Gender:**
  - What is your gender? [Male/Female/Other]
- **Age:**
  - What is your age? [Numeric input]
- **Health Self-Assessment:**
  - How do you rate your current health? [Excellent/Good/Fair/Poor/Very Poor]
- **Current Health Problems:**
  - Do you have any current health problems? [Yes/No]
- **Chronic Diseases:**
  - Do you have any chronic diseases? [Yes/No, specify if Yes]
- **Proper Nutrition:**
  - Do you follow a proper nutritional diet? [Yes/No]
- **Bad Habits:**
  - Do you have any of the following habits? [Smoking/Alcohol/None]

4. **Data Collection and Encoding:**

After collecting the survey data, we usually encode it into quantitative format.

Example Encoding in Python:

```
import pandas as pd

# Example data from the survey

data = {

    'Health_Self_Assessment': ['Excellent', 'Good', 'Fair', 'Poor'],

    'Proper_Nutrition': ['Yes', 'No', 'Yes', 'No'],

    'Bad_Habits': ['None', 'Smoking', 'Alcohol', 'None']
```

```
}


# Create a DataFrame

df = pd.DataFrame(data)


# Encode qualitative data

df['Health_Self_Assessment'] = df['Health_Self_Assessment'].map({

    'Excellent': 1, 'Good': 0.75, 'Fair': 0.5, 'Poor': 0.25, 'Very Poor': 0

})

df['Proper_Nutrition'] = df['Proper_Nutrition'].map({'Yes': 1, 'No': 0})

df['Bad_Habits'] = df['Bad_Habits'].apply(lambda x: 0 if x == 'None' else 1)


print(df)
```

5. **Visualization Using Whisker Diagrams (Boxplots):**

To visualize the impact of qualitative features on labor productivity, we are using whisker diagrams or boxplots.

Example in Python:

```
import matplotlib.pyplot as plt
import seaborn as sns

# Ensuring we have a column with labor productivity data
df['Labor_Productivity'] = [100, 80, 85, 70]  # Dummy data for illustration purposes

# Create boxplots to visualize the impact of qualitative features on labor productivity
fig, axes = plt.subplots(1, 3, figsize=(15, 5), sharey=True)

# Health Self-Assessment
sns.boxplot(ax=axes[0], x='Health_Self_Assessment', y='Labor_Productivity', data=df)
axes[0].set_title('Impact of Health Self-Assessment on LP')

# Proper Nutrition
sns.boxplot(ax=axes[1], x='Proper_Nutrition', y='Labor_Productivity', data=df)
axes[1].set_title('Impact of Proper Nutrition on LP')
```

To summarize the distribution and variation in labor productivity, descriptive statistics such as the mean, mode, median, and variation can provide a comprehensive view of data heterogeneity. Here's how to proceed:

1. **Descriptive Statistics Calculation:**

We can calculate the following statistics for the labor productivity variable:

- **Mean:** The average value of labor productivity.
- **Mode:** The most frequently occurring value.
- **Median:** The middle value when data is ordered.
- **Variance:** The measure of data dispersion.
- **Standard Deviation:** The square root of variance, showing the average distance of each data point from the mean.

2. **Presenting Descriptive Statistics in a Table:**

Here's a template for presenting the descriptive statistics in tabular form.

Example Table 2: Descriptive Statistics for Labor Productivity

Statistic	Value
Mean	85.5
Median	85
Mode	85
Variance	239.44
Standard Deviation	15.47

3. **Example Python Code to Calculate Descriptive Statistics:**

You can use Python's pandas and numpy libraries to calculate these statistics.

```
import pandas as pd
import numpy as np

# Sample data for labor productivity
data = {'Labor_Productivity': [100, 80, 85, 70, 90, 85, 95, 110, 75, 60]}

# Create a DataFrame
df = pd.DataFrame(data)

# Calculating descriptive statistics
mean_value = df['Labor_Productivity'].mean()
median_value = df['Labor_Productivity'].median()
mode_value = df['Labor_Productivity'].mode()[0]
variance_value = df['Labor_Productivity'].var()
std_dev_value = df['Labor_Productivity'].std()

# Printing the values
statistics_table = pd.DataFrame({
   'Statistic': ['Mean', 'Median', 'Mode', 'Variance', 'Standard Deviation'],
   'Value': [mean_value, median_value, mode_value, variance_value, std_dev_value]
})
```

print(statistics_table)

Output Example:

Statistic	**Value**
Mean	85.5
Median	85
Mode	85
Variance	239.44
Standard Deviation	15.47

4. **Interpreting Descriptive Statistics:**

The significant variation and non-coincidence of the mean, mode, and median indicate data heterogeneity. Here’s what these statistics suggest:

- **Mean (Average):** Indicates the central tendency of labor productivity.
- **Median (Middle Value):** Shows the midpoint of the dataset.
- **Mode (Most Frequent):** Highlights the most common value in labor productivity.
- **Variance and Standard Deviation:** Illustrate the spread and dispersion of the data.

To analyze the relationships between various factors and labor productivity, given that different measurement scales are involved, you can employ multiple statistical techniques. Here's a structured approach based on your requirements:

1. **Correlation Analysis for Continuous and Ordinal Variables:**

a. Pearson Correlation Coefficient:

Used for continuous variables like labor productivity, age, and the number of children.

b. Spearman’s Rank Correlation Coefficient:

Used for ordinal variables like education level, which provides a rank-based measure of association.

2. **Multivariate Analysis of Variance (ANOVA):**

Used to assess the influence of categorical variables on labor productivity. This includes testing hypotheses about the equality of means for the corresponding levels of categorical factors.

3. **Implementation Steps:**

a. Calculating Correlation Coefficients:

Use the Pearson correlation for continuous variables and Spearman’s rank correlation for ordinal variables.

b. Performing ANOVA:

To assess categorical variables like gender, chronic diseases, and bad habits, use ANOVA to test for differences in means.

c. 3D Surface and Contour Plots:

Visualize the relationship between labor productivity, age, and the number of children using 3D surface plots and contour plots.

4. **Example Python Code:**

a. Correlation Analysis:

```
import pandas as pd
import numpy as np
from scipy.stats import pearsonr, spearmanr
import statsmodels.api as sm
from statsmodels.formula.api import ols
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
import seaborn as sns

# Sample data
data = {
    'Labor_Productivity': [100, 80, 85, 70, 90, 85, 95, 110, 75, 60],
    'Age': [34, 28, 43, 39, 29, 30, 38, 45, 25, 22],
    'Children': [2, 1, 3, 2, 1, 1, 2, 3, 0, 0],
    'Education': [3, 2, 4, 3, 2, 2, 3, 4, 1, 1],
    'Gender': ['Male', 'Female', 'Male', 'Female', 'Male', 'Female', 'Male', 'Female', 'Male', 'Female'],
    'Chronic_Diseases': ['Yes', 'No', 'Yes', 'No', 'Yes', 'No', 'Yes', 'No', 'No', 'Yes'],
    'Bad_Habits': ['Yes', 'Yes', 'No', 'No', 'Yes', 'No', 'Yes', 'No', 'Yes', 'No']
}

# Create DataFrame
df = pd.DataFrame(data)

# Calculate Pearson correlation for continuous data
pearson_corr_age, _ = pearsonr(df['Labor_Productivity'], df['Age'])
pearson_corr_children, _ = pearsonr(df['Labor_Productivity'], df['Children'])

# Calculate Spearman correlation for ordinal data
spearman_corr_education, _ = spearmanr(df['Labor_Productivity'], df['Education'])

print(f"Pearson Correlation (Age): {pearson_corr_age}")
print(f"Pearson Correlation (Children): {pearson_corr_children}")
print(f"Spearman Correlation (Education): {spearman_corr_education}")
```

b. ANOVA for Categorical Variables:

```
# Prepare the data for ANOVA
df['Chronic_Diseases'] = df['Chronic_Diseases'].map({'Yes': 1, 'No': 0})
df['Bad_Habits'] = df['Bad_Habits'].map({'Yes': 1, 'No': 0})
df['Gender'] = df['Gender'].map({'Male': 1, 'Female': 0})

# Perform ANOVA
```

```
anova_results = sm.stats.anova_lm(ols('Labor_Productivity ~ Gender + Chronic_Diseases + Bad_Habits', data=df).fit(), typ=2)
print(anova_results)
```

c. 3D Surface and Contour Plots:

```
# 3D Surface Plot
fig = plt.figure()
ax = fig.add_subplot(111, projection='3d')

ax.scatter(df['Age'], df['Children'], df['Labor_Productivity'])

ax.set_xlabel('Age')
ax.set_ylabel('Children')
ax.set_zlabel('Labor Productivity')

plt.show()

# Contour Plot
plt.figure(figsize=(8, 6))
sns.kdeplot(x
```

REFERENCE:

1. K.H. Abdurakhmanov. Artificial Intelligence – The Foundation of Sustainable Economic Development. Monograph. Moscow (Russian Federation): Scientific Publishing House of Plekhanov Russian University of Economics. 2023. - 356 pages.
2. Abdurakhmanova G.K. The role of small businesses in the market economy //Science and practice.. - 2020. –No3 -S. 77.
3. G.K. Abdurakhmonova Human resources management Textbook T.: 2023
4. G.K. Abdurakhmonova Human resources management Textbook T.: 2020
5. A. B. Khaitov. "Human resources management": Textbook. - T.: "2019
6. Asafo-Adjei, A. B. (2007) Study on the Role of Human Resource Information Systems (HRIS) in Strategic Human Resource Management (SHRM). Flanken, Sweden
7. Beckers, A.M., Bsat, M.Z. (2002), A DSS Classification Model for Research in Human Resource Information Systems, Information Systems Management, Vol. 19 No.3, pp.41 -50.