

# მონაცემთა ანალიტიკა - Python

კურსის სახელწოდება	მონაცემთა ანალიტიკა - Python
ინსტრუქტორი	ნოდარ ოქროშიაშვილი
Email	nodar.okroshiashvili@btu.edu.ge
კურსის აღწერა	კურსი მიზნად ისახავს სტუდენტებს შეასწავლოს მონაცემთა ანალიზი პითონის საშუალებით. გააცნოს მონაცემთა ანალიზისთვის საჭირო გავრცელებული მეთოდები და ასწავლოს მათი გამოყენება რეალური ამოცანების გადასაჭრელად.
კურსის შედეგი	<p>კურსის გავლის შემდეგ სტუდენტს შეეძლება:</p> <ul style="list-style-type: none"> <li>• კომპანიის/დამსაქმებლის წინაშე მდგარი პრობლემის იდენტიფიცირება.</li> <li>• საჭირო და შესაბამისი მონაცემების მოძიება და ანალიზისათვის მომზადება.</li> <li>• მონაცემების გაწმენდა და დამუშავება. შედეგების ანალიზი.</li> <li>• მონაცემებზე დაყრდნობით კანონზომიერების შესწავლა და მისი გამოყენება გადაწყვეტილების მისაღებად.</li> <li>• მანქანური სწავლების სხვადასხვა ალგორითმების გამოყენება რეგრესიისა და კლასიფიკაციის ამოცანის გადასაჭრელად.</li> </ul>
ლექციების განრიგი	TBA
კურსის წინაპირობა	<ul style="list-style-type: none"> <li>• Python I და Python II</li> <li>• სტატისტიკა I</li> </ul>
საფასვლებული ლიგერატურა	<p>Python for Data Analysis, Wes McKinney, O'Reilly Media, Inc., 2007</p> <p>Python Data Analytics With Pandas, NumPy, and Matplotlib, Fabio Nelli, Apress Inc, 2018</p>
დამატებითი ლიგერატურა	Python Data Analysis, Ivan Idris, Packt Publishing, 2014

## კურსის შინაარსი

ლექცია	მეცადინეობის ფორმა	თემა/აქტივობა	ლიგერატურა
--------	--------------------	---------------	------------

1	ლექცია - 1სთ.	<b>შესავალი</b> <ul style="list-style-type: none"> <li>• მონაცემთა მეცნიერება</li> <li>• მონაცემთა ანალიტიკა</li> <li>• მანქანური სწავლება</li> <li>• ღრმა სწავლება</li> <li>• მონაცემთა ანალიტიკის/მეცნიერების გამოყენება</li> <li>• რატომ Python მონაცემთა ანალიზისათვის?</li> <li>• სილაბუსის გაცნობა</li> <li>• მნიშვნელოვანი ბიბლიოთეკები</li> <li>• სამუშაო გარემოს გაცნობა</li> </ul>	<ul style="list-style-type: none"> <li>• McKinney, გვ 4-8, 15-46, 51-84</li> <li>• McKinney, გვ 15-46</li> <li>• Nelli, გვ 103-121</li> </ul>
	პრაქტიკული 1სთ-1.5 სთ.	სამუშაო გარემოს მოწყობა და საჭირო ბიბლიოთეკების დაინსტალირება. მონაცემების წაკითხვა ფაილიდან და ფაილში ჩაწერა. მონაცემების ზომის და ტიპების შემოწმება, აღწერითი სტატისტიკა, უნიკალური ლაკვირებები, ლაკვირებების დათვლა.	
2	ლექცია და პრაქტიკული მეცადინეობა - 2 სთ.	<b>Introduction to Pandas</b> <ul style="list-style-type: none"> <li>• Pandas Series</li> <li>• Pandas DataFrame</li> <li>• Index, Selection, Filtering</li> <li>• Re-indexing, Dropping Entries, Axis</li> <li>• Sorting and Ranking</li> <li>• Column rename, reorder, insertion, deletion</li> <li>• Arithmetic operations</li> <li>• Statistics with DataFrame</li> </ul>	<ul style="list-style-type: none"> <li>• McKinney, გვ 123-167</li> <li>• Nelli, გვ 63-95</li> <li>• <a href="#">10 minutes to Pandas</a></li> <li>• <a href="#">Seeing Theory</a></li> </ul>
3	ლექცია და პრაქტიკული მეცადინეობა - 2 სთ.	<b>Introduction to NumPy</b> <ul style="list-style-type: none"> <li>• Scalar, Vector, Matrix, NdArray</li> <li>• Indexing and Slicing</li> <li>• Boolean Indexing</li> <li>• Universal Functions</li> <li>• Sorting</li> <li>• Statistical methods</li> <li>• Linear Algebra</li> </ul>	<ul style="list-style-type: none"> <li>• McKinney, გვ 85-122</li> <li>• Nelli, გვ 35-61</li> <li>• Idris, გვ 25-58, 59-83</li> </ul>
4	ლექცია და პრაქტიკული მეცადინეობა - 2 სთ.	<b>Data Cleaning and Preparation</b> <ul style="list-style-type: none"> <li>• Finding and filling missing values</li> <li>• Removing duplicate values</li> <li>• Replacing values</li> <li>• Discretization and Binning</li> <li>• Detecting outliers</li> <li>• String manipulations</li> </ul>	<ul style="list-style-type: none"> <li>• McKinney, გვ 191-219</li> </ul>

			<ul style="list-style-type: none"> <li>Variable Transformation</li> </ul>	
5	ლექცია პრაქტიკული მეცადინეობა 2 სთ.	ღა -	<b>Join, Combine, Reshape DataFrame</b> <ul style="list-style-type: none"> <li>Combining and Merging</li> <li>Joining and Concatenation</li> <li>Reshaping and Pivoting</li> <li>Groupby</li> <li>Cross Tabulation</li> <li>Long to Wide format</li> <li>Wide to Long format</li> </ul>	<ul style="list-style-type: none"> <li>McKinney, გვ 221-251, 287-316</li> <li>Nelli, 131-165</li> </ul>
6	ლექცია პრაქტიკული მეცადინეობა 2 სთ.	ღა -	<b>Data Visualization</b> <ul style="list-style-type: none"> <li>Line Plots</li> <li>Bar Charts</li> <li>Scatter Plots</li> <li>Histogram</li> <li>Pie Chart</li> <li>Box Plot</li> <li>The Datazaurus Dozen</li> <li>Anscombe's Quartet</li> <li>Unstructured Quartet</li> </ul>	<ul style="list-style-type: none"> <li>McKinney, გვ 253-286</li> <li>Nelli, გვ 167-235</li> <li><a href="#">Matplotlib Tutorials</a></li> <li><a href="#">PyPlot Tutorial</a></li> <li><a href="#">Example Gallery</a></li> <li><a href="#">SeaBorn</a></li> <li><a href="#">Plotly Gallery</a></li> <li><a href="#">Datazaurus Dozen</a></li> <li><a href="#">Anscombe's Quartet</a></li> </ul>
7	ლექცია პრაქტიკული მეცადინეობა 2 სთ.	ღა -	<b>Time Series</b> <ul style="list-style-type: none"> <li>Date and Time data types</li> <li>Dealing with datetime objects</li> <li>Indexing, Selection, Sub-setting</li> <li>Periods and Period Arithmetic</li> <li>Converting Timestamp to period and back</li> <li>Resampling</li> <li>Time Shifting</li> </ul>	<ul style="list-style-type: none"> <li>McKinney, გვ 317-362</li> </ul>
8	ლექცია პრაქტიკული მეცადინეობა 2 სთ.	ღა -	<b>Regression</b> <ul style="list-style-type: none"> <li>One variable regression</li> <li>Decision tree regression</li> <li>Random Forest regression</li> </ul>	<ul style="list-style-type: none"> <li>Instructor notebook</li> <li><a href="#">Ordinary Least Squares</a></li> <li><a href="#">Decision Tree Regressor</a></li> <li><a href="#">Random Forest Regressor</a></li> </ul>
9	ლექცია პრაქტიკული მეცადინეობა 2 სთ.	ღა -	<b>Classification</b> <ul style="list-style-type: none"> <li>Logistic Regression</li> <li>Decision Tree classification</li> <li>Random Forest classification</li> </ul>	<ul style="list-style-type: none"> <li>Instructor notebook</li> <li><a href="#">Logistic Regression</a></li> <li><a href="#">Decision Tree Classifier</a></li> <li><a href="#">Random Forest Classifier</a></li> </ul>
10	ლექცია პრაქტიკული მეცადინეობა 2 სთ.	ღა -	<b>Clustering</b> <ul style="list-style-type: none"> <li>Introduction to Clustering</li> <li>K-means clustering</li> <li>K-means ++</li> </ul>	<ul style="list-style-type: none"> <li>Instructor notebook</li> <li><a href="#">K-means</a></li> </ul>

11	<p>ლექცია პრაქტიკული მეტადინეობა 2 სთ.</p> <p>ლა –</p>	<p><b>Association Analysis</b></p> <ul style="list-style-type: none"> <li>• Introduction to association analysis</li> <li>• Analyzing customer basket</li> <li>• Apriori Algorithm</li> </ul>	<ul style="list-style-type: none"> <li>• Instructor notebook</li> <li>• <a href="#"><u>Market Basket Analysis</u></a></li> <li>• <a href="#"><u>Apriori Algorithm</u></a></li> </ul>
12	<p>ლექცია პრაქტიკული მეტადინეობა 2 სთ.</p> <p>ლა –</p>	<p><b>Sentiment Analysis – If time permit</b></p> <ul style="list-style-type: none"> <li>• Motivation</li> <li>• Summary of Natural Language Processing</li> <li>• Classifying positive and negative sentiments</li> </ul>	<ul style="list-style-type: none"> <li>• Instructor notebook</li> </ul>