# JOB DEMANDS-RESOURCES MODEL

*Submitted in partial fulfilment of the requirements for the degree of*

**Bachelor of Technology**
in
**Computer Science and Engineering**

*by*

**SHAURYA CHOUDHARY**
**18BCE2113**

**Under the guidance of**

**Prof. Remya L**
**School of Mechanical Engineering**
**VIT, Vellore.**


VIT®
Vellore Institute of Technology
(Deemed to be University under section 3 of UGC Act, 1956)

April, 2021

# Job Demands – Resources Model

Shaurya Choudhary
*Computer Science and Engineering*
*Vellore Institute of Technology*
Vellore, India
shaurya.src@gmail.com

***Abstract*—In this paper, we aim to study the Job Demands and Resources model and use it as a basis to analyze Job Satisfaction of employees in IT companies. For this, we use survey data from Stack Overflow and employ analytical procedures to establish findings about Job Satisfaction in different regions of the world. We conclude that the JD-R model can be applied to any industry due to its flexibility in framework and allow IT companies to establish a work engagement environment where needs of employees are given priority, which in turn yields in better performance. The proposed method determined that Europe and North America are the prime regions to achieve high job satisfaction while Women are most satisfied while working in North America. It was also found that Formal Education is not an obstacle to high job satisfaction.**

***Keywords — JD-R, Job Satisfaction, Stack Overflow, IT companies.***

## I. INTRODUCTION

In this paper, we are exploring the JD-R Model and its implications in today's industries. The Job Demands – Resources Model establishes the assumption that whereas every field of occupation may have different factors for employees' well-being, these factors can be classified into two categories: Job Demands, and Job Resources. This approach makes the JD-R Model applicable to all kinds of industries irrespective of particular demands and requirements involved.

After extensive research, it established the existence of two simultaneous processes. High job demands exhaust employees' mental and physical resources and therefore lead to the depletion of energy and to health problems. This is the health impairment process. In contrast, job resources foster employee engagement and extra-role performance. This is the motivational process. Importantly, several studies have shown that job resources may buffer the impact of job demands on stress-reactions. In addition, research has confirmed that job resources particularly have motivational potential when job demands are high.

Recent works in this domain has proved that organizational support to employees in form of Job Resources, which can be both emotional, physical, or in form of any asset, can prove to be an exponential factor in employee's motivation to fulfil his/her Job Demands. On the other hand, lack of these resources causes employees' burnout. Burnout is a special type of psychological occupational stress resulting from a response to chronic exposure to work-related stressors.

The consequences of exchanges between an organization and its employees can be both positive (e.g., engagement, organizational commitment, job satisfaction) from job resources and negative (e.g., burnout, turnover intentions) from job demands. More specifically, "interpersonal demands at works that exceed the worker's resources" can create burnout. Burnout is theoretically characterized by three components, including emotional exhaustion, depersonalization (i.e., detachment from others or indifference at work), and reduced professional efficacy (i.e., the tendency to evaluate one's efforts and achievement in a negative manner). Among these three components, emotional exhaustion is the most representative one in related studies which is found to influence the job satisfaction the most. Hence, in the following paper, we analyse the job satisfaction level of more than 64,000 developers from around 213 countries.

In this paper we make the case that the Job Demands Resources (JD-R) model can be used as an integrative conceptual framework for monitoring the workplace with the aim to increase work engagement and prevent burnout. The JD-R model is particularly suited for this purpose because: (1) it integrates a positive focus on work engagement with a negative focus on burn-out into a balanced and comprehensive approach; (2) it is having a broad scope, that allows to include all relevant job characteristics; (3) it is flexible, so that it can be tailored to the needs of any organization; (4) it acts as a common communication tool for all stakeholders. In contrast, previous models focused almost exclusively on negative aspects of the job and included a limited, predefined set of job characteristics. Because of its comprehensive, broad, flexible and communicative nature, the JD-R model not only enjoys great popularity among academic researchers, but it makes the model also quite suitable for practical use in organizations.

This paper starts with a brief description of the JD-R model and goes through a literature review of recent

research journals and then introduces an extensive job satisfaction analysis of employees in tech related industries from all around the world, and assess the effect of gender on job satisfaction in a broader scale of work environment across different world regions and also the contribution of formal education. The analysis is supported by appropriate data visualizations followed by overall discussion of the work.

## II. OBJECTIVES

The objective of this project is to understand the concepts of Job Demands-Resources model and its implications. In the initial stage of this work, the problem statement is identified by going through several recent Research Journals in the JD-RM domain. We further perform data analysis on a survey dataset to identify the job satisfaction level of Engineers from all around the world, and extract others factors which contribute to the satisfaction level of employees in the working environment.

The main objectives we try to achieve are:

*1)* Review of recent Research journals on JD-R Model.

*2)* Extracting appropriate dataset for analysis.

*3)* Analyzing the Job Satisfaction level of employees.

*4)* Job Satisfaction distribution among the genders.

*5)* Influence of Formal Education on Job Satisfaction.

## III. RELATED WORK

For identification of background and recent trends in the domain of Job Demands-Resources Model, we go through five recent research journals covering different aspects and implications of the JD-R model. The literature study provided some great insights in exchange between an organization and its employees:

- Employees' satisfaction exerts a positive and significant effect on employees' productivity.
- If individuals believe that their organization exerts significant effort to increase their happiness and satisfaction level at work, they will feel morally obliged to reciprocate by showing positive attitudes and behavior that can enhance and advance the productivity of their organization.
- Job resources are positively related to work engagement and enhance employee well-being.

The above inferences were drawn after studying all of the five journals taken for related work of this paper. Each referenced journal is discussed below in brief as the part of literature review.

### *A. JD-R Model Amd Workplace Internet Leisure*

The main aim of this study is to explore about Workplace Internet Leisure Policy and its' role in integration with JD-R model. It examines the effect of WIL on workplace outcome variables such as employee satisfaction (ES) and employee productivity in the Malaysian context. For studying the hypothesis, the research paper uses Partial least square technique using SmartPLS-3 on collected data of 282 respondents.

Findings reveal that workplace WIL, workplace WIL policy and workplace autonomy orientation (WAO) affect employees' satisfaction. Additionally, the mediating role of ES was found to be significant. This is a pioneering study which suggests that moderate use of WIL can have a positive and significant effect on workplace outcome variables. Moreover, this study theorized ES as a mediating variable; this helps to explain how organizations can transform workplace resources in term of internet leisure, WILP and WAO into high productivity by elevating employees' satisfaction. [1]

### *B. Working Conditions Prediction Using JD-R Model*

The journal applies the JD-R model and a new categorization approach to study the relationship between working conditions and innovation. By applying confirmatory factor analysis and structural equation modelling to a cross-sectional online study, it showed that two types of demands, hindrance and challenge, and two types of job resources, task-related and social, represent different types of working conditions with respect to innovation.

Task-related and social job resources positively predicted individual innovation. Social job resources and challenge job demands revealed a positive association with perception of organizational innovation, whereas hindrance job demands were negatively related to it. The relevance of the studied types of working conditions for individual and perceived organizational innovation varied. [2]

### *C. Test of JD-R Model in Thailand*

The purpose of this paper is to test the hypothesized relationships between job demands, job resources and personal resources toward work engagement, by utilizing a cross section of Thai employees. In this study, a group of 416 employees responded to a set of

self-report surveys on job demands, job resources, personal resources and work engagement.

The results of the hierarchical regression analysis supported the relationships between job demands (i.e., workload and role conflict), job resources, personal resources (self-efficacy) and work engagement. In addition, the results supported the role of (positive) self-esteem as moderator in the role ambiguity and work engagement relationship, and the role of self-efficacy in buffering the effect of role conflict and workload on work engagement. The final model explained 43 percent of the variance in the dependent variable. [3]

### *D. Emotional Exhaustion and Well-Being Analysis*

The paper is focused on Taiwanese Bus Drivers as compared with other occupations, urban bus drivers work under conditions that are among the most demanding, stressful, and unhealthy conditions with higher rates of mortality and morbidity. Bus drivers must balance various requests from passengers and management, as well as follow traffic rules and regulations when they are on duty. Hence, these work-related expectations and requirements increase bus drivers' workload and stress, subsequently leading to their job burnout and a negative impact on their well-being. In accordance with the job demand-resource (JD-R) model, job demands and job resources, two specific risk factors of every occupation, can explain employees' well-being.

Burnout is theoretically characterized by three components, including emotional exhaustion, depersonalization, and reduced professional efficacy. This study uses emotional exhaustion to represent individual's burnout. The results show that emotional exhaustion does play an important role in the work domain. It identifies role overload and work-family conflict as two stressors related to job demands and organizational support as the job resource factor to affect emotional exhaustion in bus driver context. Results reveal that only organizational support instead of supervisor support has a hindering effect on emotional exhaustion. It also implies that drivers' perception of whether the bus company is concerned about them and cares about their needs and welfare plays an important role in lessening the occurrence of emotional exhaustion. [4]

### *E. Influence of Job and Individual Resources*

This journal aims to explore the drivers of work engagement through perceived organizational support and regulatory emotional self-efficacy among Chinese police officers using a convenient sampling method. The study considers a sample population of 744 Chinese Police Officers. It assumes that a moderated mediation model, in which job satisfaction plays a mediating role in the relationship between perceived organizational support and work engagement regulatory emotional self-efficacy moderates not only the relationship between perceived organizational support and job satisfaction but also the relationship between job satisfaction and work engagement.

Job satisfaction mediated a positive relationship between organizational support and work engagement, and the perceived organizational support-job satisfaction and the job satisfaction-work engagement relationships were positively moderated by regulatory emotional self-efficacy, such that these relationships were stronger at higher levels of regulatory emotional self-efficacy. These findings have a practical significance for Chinese police officers' work engagement advancement. [5]

## IV. PROPOSED WORK

### *A. Dataset Description*

For implementation of this project, we have selected Stack Overflow Annual Developer Survey Dataset. The dataset is prepared annually by Stack Overflow after conducting a massive survey of more than 64,000 developers from around 213 countries. They survey all sorts of information like programming languages, salary, code style and various other information.

The dataset originally consists of two files: survey_results_public, and survey_results_schema. The public file contains the main survey results with one respondent per row and one column per answer whereas the schema file contains the questions that correspond to each column name.

### *B. Data Pre-Processing*

Data pre-processing is a crucial stage for performing any analysis to test a hypothesis. For the scope of this project, the main aim during our data pre-processing would be to format the data according to our needs by dropping irrelevant attributes and also grouping the data into categories for analysis.

Initially, we group all the available countries in the dataset according to their world region by generating a new dictionary csv file which maps country to their respective world regions. Followed by this, we identify the different features of the dataset and extract the null values contribution in each column. At last, we use LabelEncoder to encode numerical values to the categorical dataset attributes.

The dataset is updated after performing the above pre-processing methods and used further to perform data analysis and visualization.

### C. *Modules Description*

The project is divided into 3 main modules each addressing a different agenda for analysis.

a) *Job Satisfaction around World:* In the first module, we assess the average job satisfaction of employees in different countries and group them according to the World Region. The module identifies in which region of world, the developer employees are more satisfied.

b) *Influence of Gender on Job Satisfaction:* This part of analysis aims to find out if both Men and Women feel the same way about their Jobs across the world. We further classify the results according to the Job Seeking Status of the employees.

c) *Role of Formal Education:* At last, we assess the influential role of formal education in job satisfaction of the developers. We also analyse the educational demands in different regions of the world and respective satisfaction level of developers in the region.

### D. *Software Requirements*

The project is implemented on Jupyter Notebook using Python 3.x. Various python libraries are used during the development of this project which must be imported/installed to run the notebook:

- sklearn
- csv
- numpy
- pandas
- matplotlib
- seaborn

The above libraries serve different purposes in the project like data pre-processing, visual graphs, data-frame manipulation, high-level mathematical operations, and dataset extraction.

### E. *Implementation*

The complete implementation of this project can be found on my GitHub repository: https://github.com/shaurya-src/job-satisfaction-analysis.

## V. RESULTS AND DISCUSSION

In today's global markets anyone could work anywhere they would like. But the average job satisfaction of an employee varies greatly in different regions of the world. The inferences drawn from the proposed work of this paper, it provides a deep insight about the satisfaction of employees or developers in IT companies.

The survey dataset consisted of more than 150 questions spanning over 19000 reviews from 165 countries and territories. It aimed to understand a number of aspects of jobs related to software development and data analytics. The analysis of the dataset resulted in the findings which are discussed below.

The vast majority of respondents were from the U.S.A when looked at by individual country. However, when grouping the respondents by world region rather than investigating by country we can see that respondents come mainly from 2 Regions: North America, and Europe, which represent about 30% of the respondents each. This was followed bs 2 regions: Eastern Europe, and South-Central Asia, which represent about 10% of the respondents each. The above four regions constituted more than 80% of all respondents in the dataset. Therefore, more emphasis has been given to these 4 world regions in further analysis.

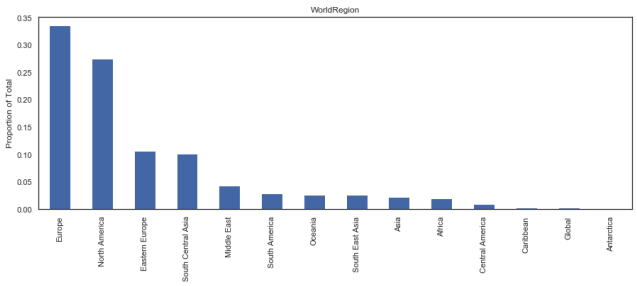



**Fig. 1. Distribution of respondents across different world region**

The proposed methodology not only focused on the job satisfaction rating from the dataset but also related it to the job-seeking status of the respondent. The job satisfaction was measured on scale of 1 to 10, with 10 being the highest Job Satisfaction of an employee.

| WorldRegion | JobSatisfaction |
|---|---|
| Eastern Europe | 6.896288 |
| Europe | 6.813744 |
| North America | 6.893957 |
| South Central Asia | 5.980796 |

**Fig. 2. Average job satisfaction**

The job seeking proportion of respondents was grouped by the world regions and it was observed that most of the respondents were settled in their current position, but were open to new opportunities.

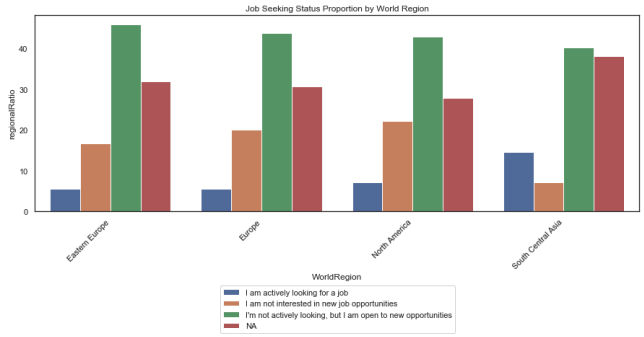



**Fig. 3. Proportions of different job-seeking status**

On merging the two results, we get the job satisfaction of employees in different regions of world categorized by job-seeking status.

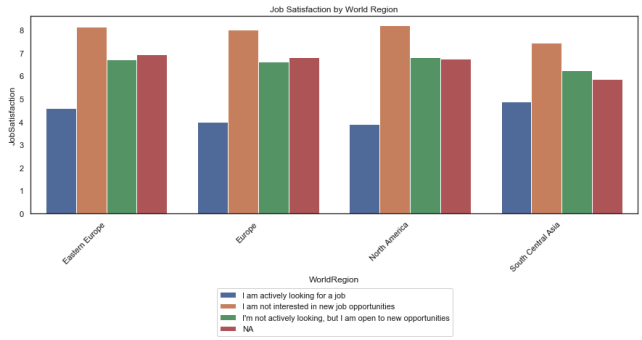



**Fig. 4. Job satisfaction categorized by job-seeking status**

The findings provide a solid case that if somebody is trying to decide to choose the best world region to work in, Eastern Europe does seem to be an attractive option. If that is not a consideration, a logical choice with consideration of job satisfaction in mind will be between Europe and North America.

The findings of next proposed module addressed if women and men feel the same way about their jobs across world regions. It was evident that women have an edge on job satisfaction in North America followed by Eastern Europe and Europe.

| WorldRegion | Gender | JobSatisfaction |
|---|---|---|
| Eastern Europe | Female | 6.977011 |
| | Male | 6.999096 |
| Europe | Female | 6.710084 |
| | Male | 6.932485 |
| North America | Female | 7.105590 |
| | Male | 6.967413 |
| South Central Asia | Female | 5.566667 |
| | Male | 6.188302 |

**Fig. 5. Job Satisfaction classified by gender**

Further analysis provided more deeper insights.

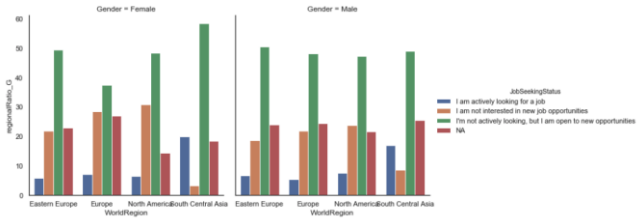



**Fig. 6. Gender distribution a/c to job status**

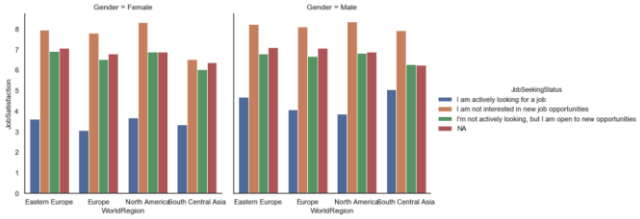



**Fig. 7. Job satisfaction by gender and job status**

At a high-level job satisfaction for women is lower than their male counterparts, except for North America where women seem to be happier with their jobs. For Women North America is the best place to be. Higher job satisfaction, more opportunities and less desire to move around are indications of a better professional life for women in this region.

The final module assessed whether the formal education plays any role later in the job satisfaction of employees in their work engagement. Overall, it was found that people with a Doctoral degree or those who have never completed a formal education are the most satisfied in their respective jobs.

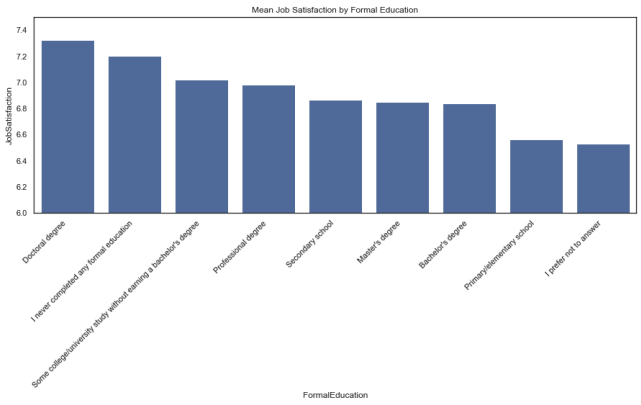



**Fig. 8. Job satisfaction at different levels of Formal Education**

The first inference appeared to still hold true for North America and Europe when categorized by world region. Professionals in these two regions have good chances at a good job even if they have no formal college degree.

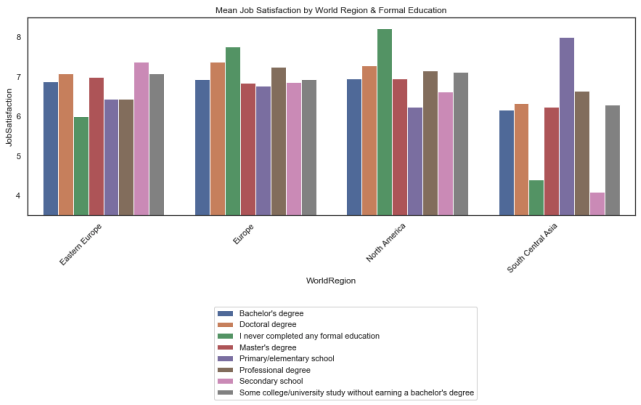



**Fig. 9. Influence of Formal Education on Job satisfaction across world**

Europe has better job satisfaction across education levels, indicating there are good jobs for professionals from all walks of life. In North America however, professionals with only primary or high school level education are not as happy.

Europe is the place to be for professionals at all levels of education. In North America at least some college or a professional degree would be required to have a satisfying job although those without completing a formal education are the most satisfied.

## VI. Future Work

Despite the insights and contributions of this project, there are limitations that need to be addressed and considered for future research. The collected data refers to the employees of tech related companies from all around the world, and not to specific departments of those companies. Therefore, in future research, data should be separated by departments to perceive the most meaningful variables of job satisfaction of each IT company department. It would be interesting to understand the reasons that make top managers satisfied with their position at an IT company and which factors have the most relevance to their satisfaction. Further, a more recent survey dataset may provide better insight in the shift of job satisfaction matrix after the Covid-19 Pandemic situation.

## VII. Conclusion

In this article, it is argued and illustrated that the JD-R model may serve as the guiding principle for an organizational development process that aims to increase work engagement and prevent burnout. The model is well equipped for this purpose because it is comprehensive, as it includes both a positive motivational process as well as a negative a stress process. This balanced approach is an important asset to ‘sell’ the model to organizations because it integrates an occupational health approach (reducing job stress and burnout) with an HR-approach (increasing work motivation and engagement).

We analyse how satisfied Professional Developers or Professionals with Coding Experience are with their jobs across regions of the world and draw inferences from the same after applying the above proposed work on the dataset.

The analyses are based on the world regions 80% of respondent reside in.

i. It is determined that Europe and North America are the prime regions to achieve high job satisfaction with Eastern Europe coming out slightly ahead.

ii. For Women the best place for a satisfying job is North America.

iii. Professionals with doctoral degrees or who have not completed formal education are most satisfied, with Europe providing good opportunities for all. Meaning Formal Education is not an obstacle to high job satisfaction.

Employees are a valuable resource for an IT company to survive and thrive. An IT company needs their employees to feel satisfied to achieve the overall objectives and to remain loyal to the company in order to achieve company success. The employees' satisfaction is the premise of this commitment and dedication. Providing the necessary conditions for an employee to feel satisfied, employees can become a priceless asset. They can contribute in so many ways for a company to achieve competitive advantage in a globalized world. Employee satisfaction can help in reducing turnover, which is high in the dynamic IT sector.

The achieved findings characterize the satisfaction of level of employees from all around the world and identifies the influence of gender and formal education in the same. The obtained results can be used as a basis for identifying the shortcomings in the working culture and environment in different regions of the world. IT industries could then improve their work engagement approach through the use of JD-R model as the work of a satisfied employee can benefit both employees and IT companies.

## ACKNOWLEDGMENT

This paper and the research behind it would not have been possible without the exceptional support of my faculty, Prof. Remya L. Her enthusiasm, knowledge and exacting attention to detail have been an inspiration and kept my work on track. I am also grateful for the insightful comments offered by my peer friends. The generosity and expertise of one and all have improved this study in innumerable ways and saved me from many errors; those that inevitably remain are entirely my own responsibility.

Finally, it is with true pleasure that I acknowledge the support of my family throughout the completion of this paper.

## REFERENCES

[1] Mohammad, J., Quoquab, F., Halimah, S., & Thurasamy, R. (2019). Workplace internet leisure and employees' productivity. Internet Research.

[2] Adler, M., & Koch, A. K. (2017). Expanding the job demands-resources model to classify innovation-predicting working conditions. management revue, 175-203.

[3] Kunte, M., & Rungruang, P. (2019). Test of the job demand resources model in Thailand. International Journal of Organization Theory & Behavior.

[4] Chen, C. F., & Hsu, Y. C. (2020). Taking a Closer Look at Bus Driver Emotional Exhaustion and Well-Being: Evidence from Taiwanese Urban Bus Drivers. Safety and Health at Work, 11(3), 353-360.

[5] Lan, T., Chen, M., Zeng, X., & Liu, T. (2020). The influence of job and individual resources on work engagement among Chinese police officers: a moderated mediation model. Frontiers in psychology, 11.

# JOB SATISFACTION ANALYSIS

[ Job Demands-Resources Model ]

Shaurya Choudhary - 18BCE2113

## Data Pre-Processing

In [4]:
```
import csv
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.preprocessing import LabelEncoder

%matplotlib inline

df = pd.read_csv('../Data/survey_results_public.csv')
schema = pd.read_csv('../Data/survey_results_schema.csv')
reader = csv.reader(open('../Data/countryRegionDic.csv'))
df.head()
```

Out[4]:

| | Respondent | Professional | ProgramHobby | Country | University | EmploymentStatus | FormalEduc |
|---|---|---|---|---|---|---|---|
| **0** | 1 | Student | Yes, both | United States | No | Not employed, and not looking for work | Secondary s |
| **1** | 2 | Student | Yes, both | United Kingdom | Yes, full-time | Employed part-time | college/univ study w earr |
| **2** | 3 | Professional developer | Yes, both | United Kingdom | No | Employed full-time | Bachelor's d |
| **3** | 4 | Professional non-developer who sometimes write... | Yes, both | United States | No | Employed full-time | Doctoral d |
| **4** | 5 | Professional developer | Yes, I program as a hobby | Switzerland | No | Employed full-time | Master's d |

5 rows × 154 columns

In [5]:
```
# Displaying Column Descriptions of Dataset

styles = [dict(selector = "th", props = [('text-align', 'center')])]

with pd.option_context('display.max_colwidth', 300):
    display(schema.head().style.set_properties(**{'text-align': 'left'}).set_table_s
```

| | Column | Question |
|---|---|---|
| 0 | Respondent | Respondent ID number |
| 1 | Professional | Which of the following best describes you? |
| 2 | ProgramHobby | Do you program as a hobby or contribute to open source projects? |
| 3 | Country | In which country do you currently live? |
| 4 | University | Are you currently enrolled in a formal, degree-granting college or university program? |

In [6]:
```
#This code was used to help generate the country to world region dictionary csv file
# Sending Schema to an Excel file for closer inspection
#schema.to_excel('schema.xlsx')

#Sending column Country to an excel file for closer inspection
#df['Country'].to_excel('country.xlsx')
```

In [7]:
```
num_rows = df.shape[0] #Provide the number of rows in the dataset
num_cols = df.shape[1] #Provide the number of cols in the dataset
num_countries = df['Country'].nunique() #provide number of countries spanning the su

print('number of rows: ',num_rows)
print('number of columns: ',num_cols)
print('Number of countries:' , num_countries)
```

```
number of rows:  19102
number of columns:  154
Number of countries: 165
```

In [8]:
```
#Provide a set of columns with 0 missing values.
no_nulls = set(df.columns[df.isnull().mean()==0])

print(no_nulls)
```

```
{'ProgramHobby', 'Professional', 'EmploymentStatus', 'Country', 'University', 'Respo
ndent', 'FormalEducation'}
```

In [9]:
```
#Provide a set of columns with less than 25% missing values but not 0 missing.
twentyfive_zero_nulls = set(df.columns[df.isnull().mean() <= 0.25])
twentyfive_nulls = twentyfive_zero_nulls - no_nulls

print(twentyfive_nulls)
```

```
{'MajorUndergrad', 'YearsProgram', 'YearsCodedJob', 'CareerSatisfaction', 'CompanySi
ze', 'ClickyKeys', 'CompanyType', 'HomeRemote', 'PronounceGIF', 'JobSatisfaction'}
```

In [10]:
```
#Provide a set of columns with all missing values.
all_nulls = set(df.columns[df.isnull().mean() == 1])

print(all_nulls)
```

```
set()
```

In [11]:
```
#Provide a set of columns with 75% + missing values.
most_nulls = set(df.columns[df.isnull().mean() > 0.75])

print(most_nulls)
```

```
{'MobileDeveloperType', 'ExCoderActive', 'NonDeveloperType', 'ExCoderWillNotCode',
'ExCoderBalance', 'ExCoderBelonged', 'ExCoderNotForMe', 'ExCoderSkills', 'TimeAfterB
ootcamp', 'WebDeveloperType', 'ExCoderReturn', 'ExCoder10Years', 'YearsCodedJobPas
t', 'ExpectedSalary'}
```

## Helper Functions

In [12]:
```
def get_description(column_name, schema=schema):
    '''
    INPUT - schema - pandas dataframe with the schema of the developers survey
            column_name - string - the name of the column you would like to know abo
    OUTPUT -
            desc - string - the description of the column
    '''
    desc = list(schema[schema['Column'] == column_name]['Question'])[0]

    print(desc)

    return

def get_survey_vals(column_name, num_vals, dataset = df):
    '''
    Provide the first values of a specific column i the dataset

    INPUT: - A string for the column name of interest
           - Integer for the number of values of interest
           - the dataset of interest

    OUTPUT: - the first num_vals values of the column of interest
    '''
    col = dataset[column_name].head(num_vals)

    print(col)

    return
```

In [13]:
```
get_description('JobSatisfaction');
get_survey_vals('JobSatisfaction', 20);
```

```
Job satisfaction rating
0     NaN
1     NaN
2     9.0
3     3.0
4     8.0
5     NaN
6     6.0
7     7.0
8     6.0
9     8.0
10    9.0
11    6.0
12    NaN
13    6.0
14    8.0
15    8.0
16    NaN
17    8.0
18    3.0
19    8.0
Name: JobSatisfaction, dtype: float64
```

In [14]:
```
"""
Creating a dictionary to assign a world region to a country.
dictionary is created from a CSV file containing all countries
in the Stackoverflow dataset to which a world region has been assigned.
"""


country2Region_dic={}
for row in reader:
    country2Region_dic[row[0]]=row[1]
```

```
##print(country2Region_dic)

"""
Adding a column to the dataset dataframe defining world region from the country colu
"""

# Mapping the dictionary keys to the data frame.
df['WorldRegion'] = df['Country'].map(country2Region_dic)
df.head()
```

Out[14]:

| | Respondent | Professional | ProgramHobby | Country | University | EmploymentStatus | FormalEduc |
|---|---|---|---|---|---|---|---|
| **0** | 1 | Student | Yes, both | United States | No | Not employed, and not looking for work | Secondary s |
| **1** | 2 | Student | Yes, both | United Kingdom | Yes, full-time | Employed part-time | college/univ study w earr |
| **2** | 3 | Professional developer | Yes, both | United Kingdom | No | Employed full-time | Bachelor's d |
| **3** | 4 | Professional non-developer who sometimes write... | Yes, both | United States | No | Employed full-time | Doctoral d |
| **4** | 5 | Professional developer | Yes, I program as a hobby | Switzerland | No | Employed full-time | Master's d |

5 rows × 155 columns

In [15]:

```
def get_world_region(country, dic = country2Region_dic):
    """
    This function looks up the world region for a given country

    INPUT: - country: a text string
           - a dictionary mapping the world region to each country in the dataset
             the dictionary is created fomr an imported csv delimite file './country
    OUTPUT: - prints the World region a country is in
    """

    print(dic[country])

    return


def get_regions():
    '''
    This function retrieves all the regions represented in the dataset

    INPUT: - a dictionary mapping the world region to each country in the dataset
             the dictionary is created fomr an imported csv delimite file './country

    OUTPUT: - list of World Regions represented in the dataset assigned via the dict
```

```
    '''

    print(sorted(list(set(list(country2Region_dic.values())[1:]))))

    return

def get_countries(world_region):
    ''' This function retrieves all of the countries represented in a world region

    INPUT: - world_region:  a string of the world region of interest
           - a dictionary mapping the world region to each country in the dataset
             the dictionary is created fomr an imported csv delimite file './country

    OUTPUT: - A list of countries in the world region of interest
    '''

    for country, region in country2Region_dic.items():
        if region == world_region:
            print(country)

    return

get_regions()
get_world_region('China')
get_countries('South Central Asia')
```

```
['Africa', 'Antarctica', 'Asia', 'Caribbean', 'Central America', 'Eastern Europe',
'Europe', 'Global', 'Middle East', 'North America', 'Oceania', 'South America', 'Sou
th Central Asia', 'South East Asia']
Asia
Afghanistan
Bangladesh
Bhutan
India
Kazakhstan
Kyrgyzstan
Nepal
Pakistan
Sri Lanka
Tadjikistan
Uzbekistan
```

# MODULE 1: VARIATION OF JOB SATISFACTION BY WORLD REGION

Below we can see that the vast majority of respondents is from the U.S.A when looked at by individual country. However when grouping the respondents by world region rather than investigaintg by country we can see that respondents come mainly from 2 Regions, which represent about 30% of the respondents each:

- North America
- Europe

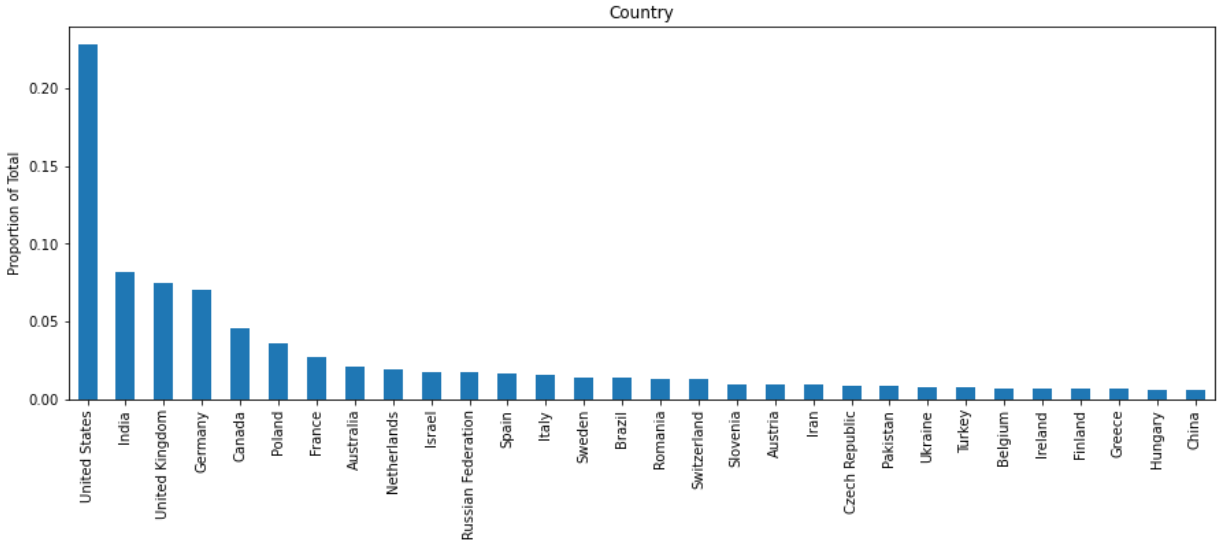
This is followed at some distance by:

- Eastern Europe
- South Central Asia

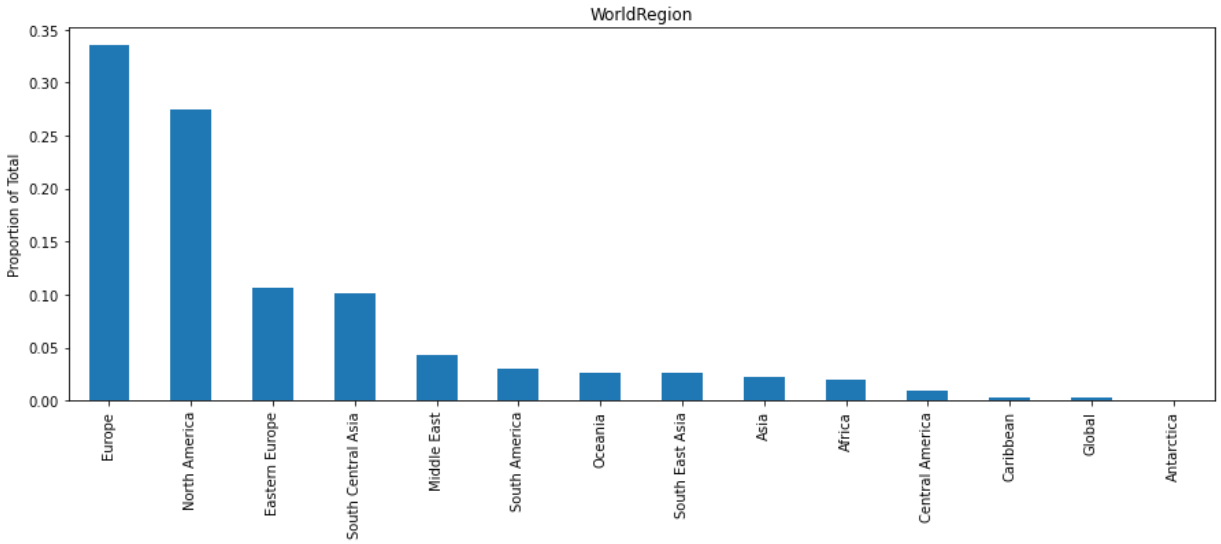
which represent approximately 10% of respondents each.

In total, the above world regions represent about 80 % of all repondents to the survey. Therefore I will continue my analyses for only these 4 regions and furthermore i will be mostly concentrating on North America and Europe.

In [16]:
```
#Provide a pandas series of the counts for country represented
count_vals = df.Country.value_counts()

# print(count_vals)

# below a bar chart of the proportion of individuals in each country for the top 30
plt.figure(figsize=(15,5))
(count_vals[:30]/df.shape[0]).plot(kind="bar");
plt.ylabel("Proportion of Total")
plt.title("Country");
```



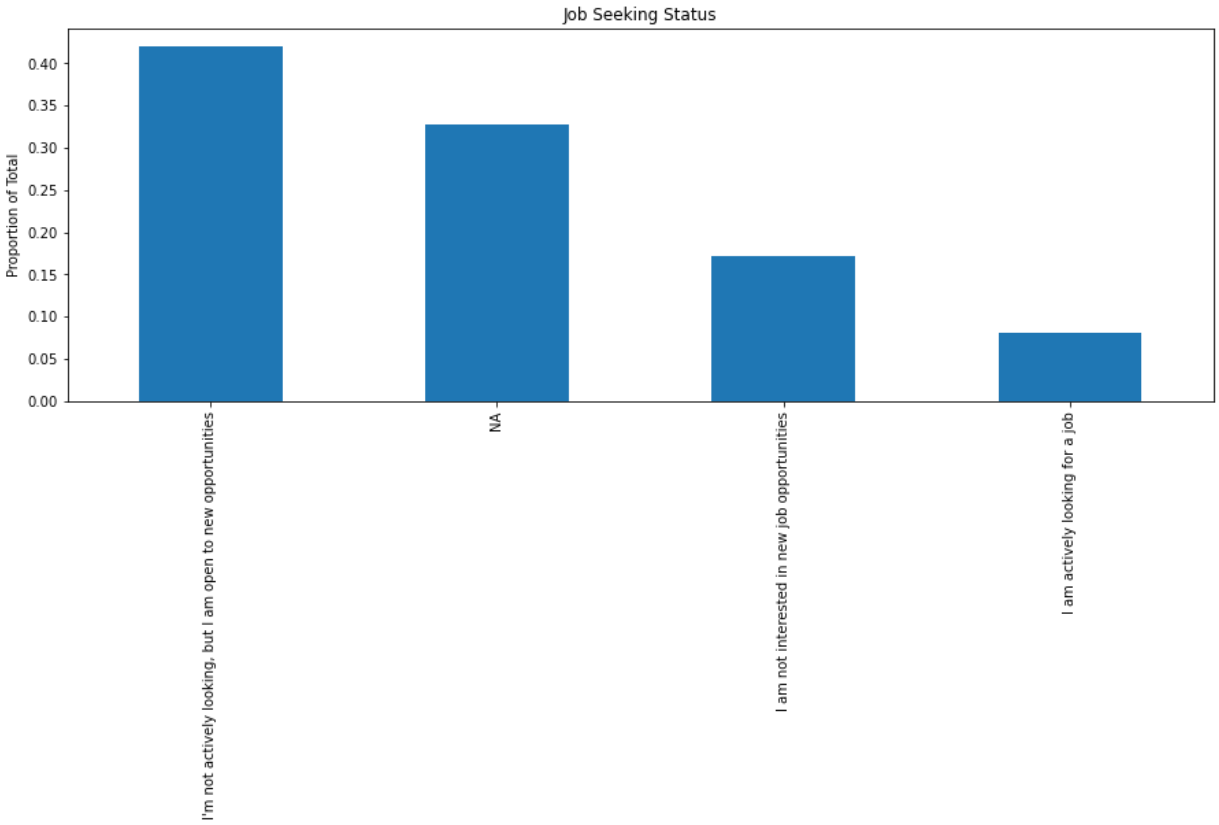

In [17]:
```
#Provide a pandas series of the counts for each World Region
count_vals = df.WorldRegion.value_counts()

# below a bar chart of the proportion of individuals in each World Region
plt.figure(figsize=(15,5))
(count_vals/df.shape[0]).plot(kind="bar");
plt.ylabel("Proportion of Total")
plt.title("WorldRegion");
```



When looking for job satisfaction I will be looking at the 'JobSatisfaction' and the 'JobSeekingStatus' columns in the dataset.

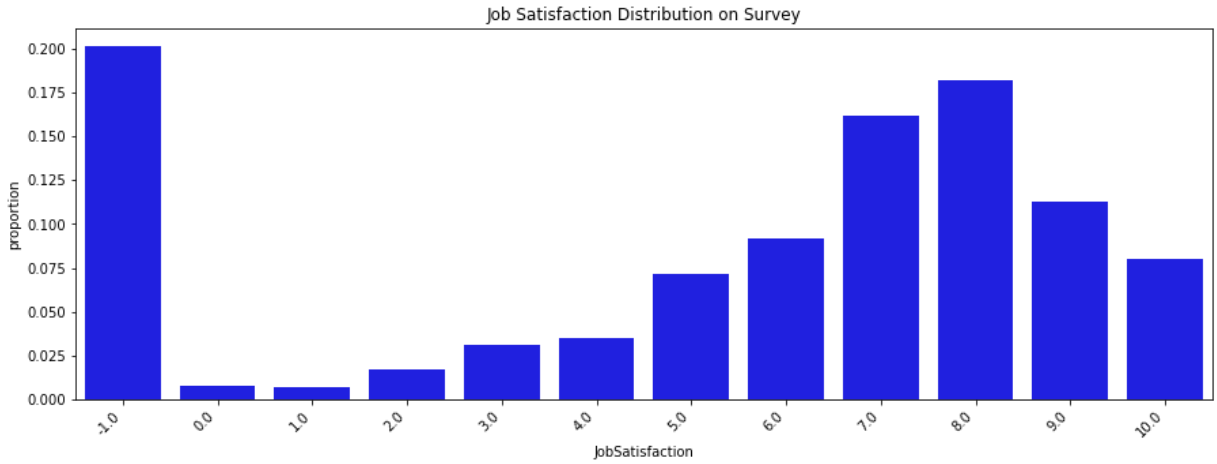
Furthermore, over 80% are professionals, either developers or non-developers with some coding experience. For this reason I will be concentrating on them and not include students or others.

In [18]:
```
#Provide a pandas series of the counts for each Job Seeking Status
df['JobSeekingStatus'].fillna(value = 'NA', inplace = True)
count_vals = df.JobSeekingStatus.value_counts()

# below a bar chart of the proportion of individuals for each Job Seeking Status
plt.figure(figsize=(15,5))
(count_vals/df.shape[0]).plot(kind="bar")
plt.ylabel("Proportion of Total")
plt.title("Job Seeking Status");
```




In [19]:
```
#Provide a pandas series of the counts for each Job Satisfaction value

df['JobSatisfaction'].fillna(value = -1, inplace = True) # replacing Nans with -1
#df['JobSatisfaction'].apply(pd.to_numeric)
df_sorted = df.sort_values('JobSatisfaction', ascending=True) #.reset_index(inplace=
count_vals = df_sorted.JobSatisfaction.value_counts()
df_count_vals = pd.DataFrame(
                data=count_vals,
                )
df_count_vals.reset_index(level=0, inplace=True)
df_count_vals.columns = ['JobSatisfaction','Count']
df_count_vals['proportion'] = df_count_vals['Count']/df.shape[0]

# below a bar chart of the proportion of individuals for each Job Seeking Status

#(count_vals/df.shape[0]).plot(kind="bar");
#plt.title("Job Satisfaction");

plt.figure(figsize=(15,5))
```

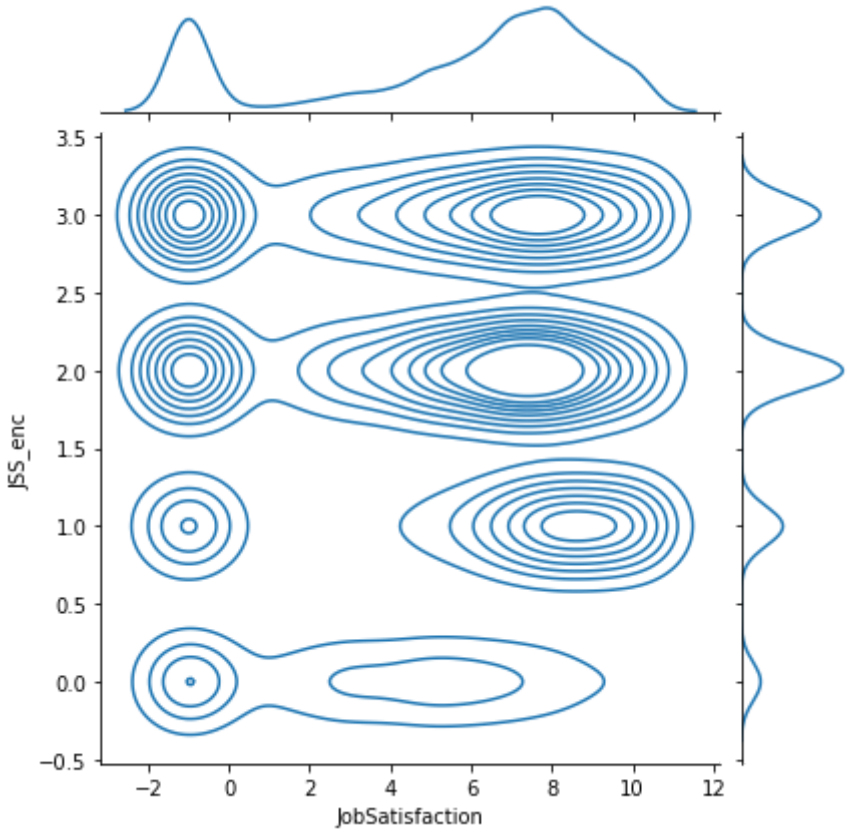
```
chart = sns.barplot(data = df_count_vals,
                    x = 'JobSatisfaction',
                    y = 'proportion',
                    #hue = 'WorldRegion',
                    ci = None,
                    color = 'b')
chart.set_title('Job Satisfaction Distribution on Survey')
chart.set_xticklabels(chart.get_xticklabels(), rotation=45, horizontalalignment='rig
#chart.legend(loc='center left', bbox_to_anchor=(1.25, 0.0), ncol=1);
```




In [20]:

```
# need to encode the categorical values of JobSeekingStatus to a numerical value put

labelencoder= LabelEncoder() #initializing an object of class LabelEncoder
#df = df[['Product_Code', 'Country_Code']].apply(le.fit_transform)

df['Gender'].fillna(value = 'NA', inplace = True)

df['JSS_enc'] = df[['JobSeekingStatus']].apply(labelencoder.fit_transform)
df['G_enc'] = df[['Gender']].apply(labelencoder.fit_transform)
df_enc = df

#Filtering out the non response for job satisfaction
#df_enc = df_enc[df_enc['JobSatisfaction'] != -1.0]

# This prints the mapping between the JobSeeking Status and the encoded value JSS_en
print(df_enc[['JobSeekingStatus','JSS_enc']].groupby('JobSeekingStatus').apply(lambd


plt.figure(figsize=(15,5));
chart = sns.jointplot(data = df_enc,
            x = 'JobSatisfaction',
            y = 'JSS_enc',
            #hue = 'JobSeekingStatus',
            #ci = None,
            kind = 'kde');
#chart.set_title('Job Seeking Status vs. Job Satisfaction')
#chart.set_xticklabels(chart.get_xticklabels(), rotation=45, horizontalalignment='ri
#chart.legend(loc='center left', bbox_to_anchor=(0.3, -0.5), ncol=1);
```

```
JobSeekingStatus
I am actively looking for a job                                 [0]
I am not interested in new job opportunities                    [1]
I'm not actively looking, but I am open to new opportunities    [2]
NA                                                              [3]
dtype: object
<Figure size 1080x360 with 0 Axes>
```




The above plot shows us the correlaton and spread between job satisfaction and job seeking status.

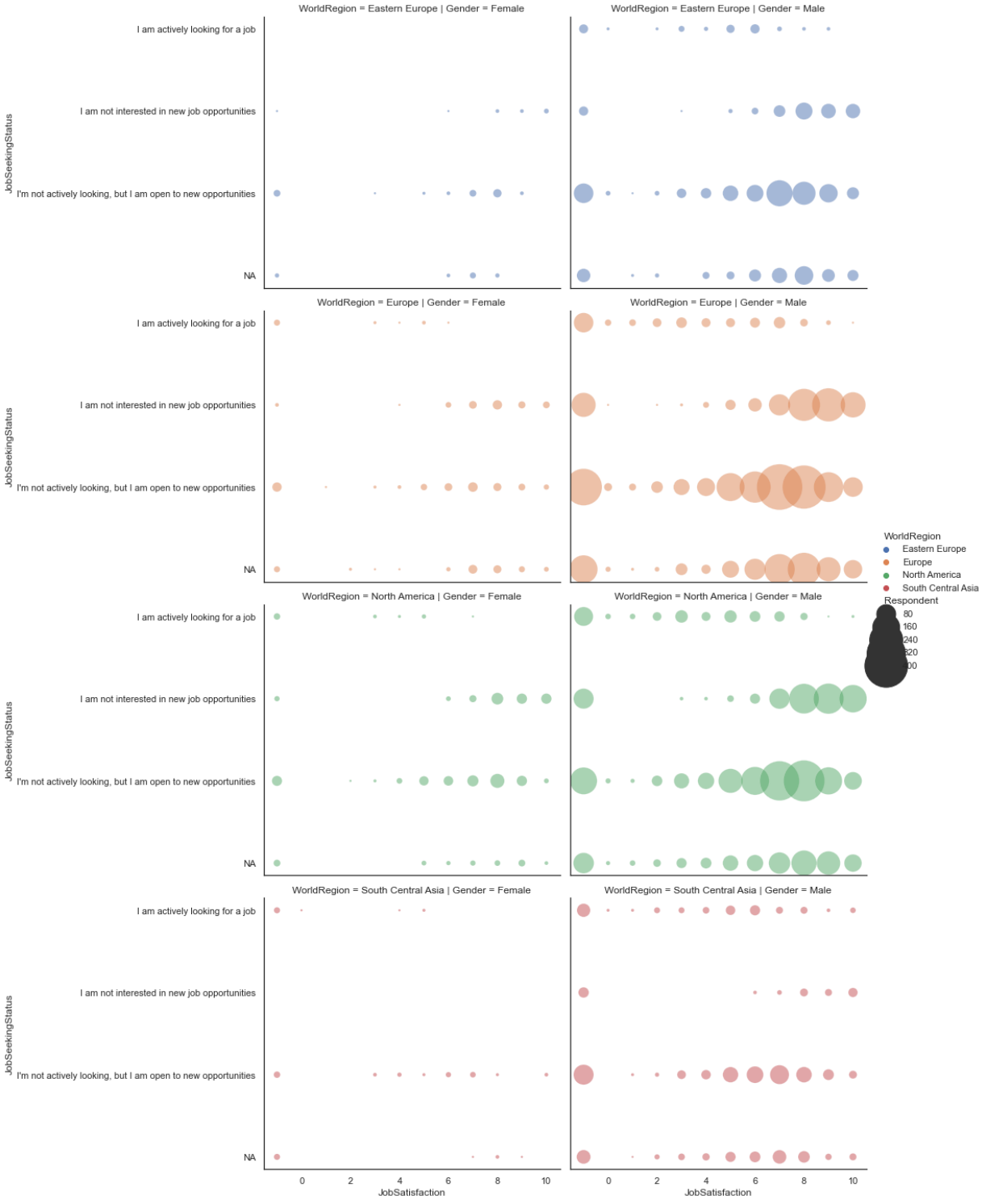
So, if we concentrate on the right side of the plot we can see a low job satisfaction relating to people actively looking for a job (0 on JSS_enc) and a high job satisfaction relating to people not looking for a job at all (1 on JSS_enc). A high number of respondents with Job Satisfaction around 7.5 are very open to new opportunities but are not actively seeking (2 on JSS_enc)

In [21]:
```
'''
Filter dataset on:
    - Gender
    - WorldRegion : - Eastern Europe
                    - Europe
                    - North America and
                    - South Central Asia
'''

df_enc = df_enc[df_enc.Gender.isin(['Male',
                                    'Female'])]
df_enc = df_enc[df_enc.WorldRegion.isin(['Eastern Europe',
                                         'Europe',
                                         'North America',
                                         'South Central Asia'])]

# Grouping Dataset to get the count per occurence
count = df_enc.groupby(['WorldRegion',
                        'Gender',
                        'JobSeekingStatus',
                        'JobSatisfaction',
                        ]).count()

# resetting index
count.reset_index(inplace=True)
```

```
# adding column representing percentage of counts of above group by over total occur
count['resp_ratio'] = 100 * count['Respondent']/count['Respondent'].sum()
```

In [22]:

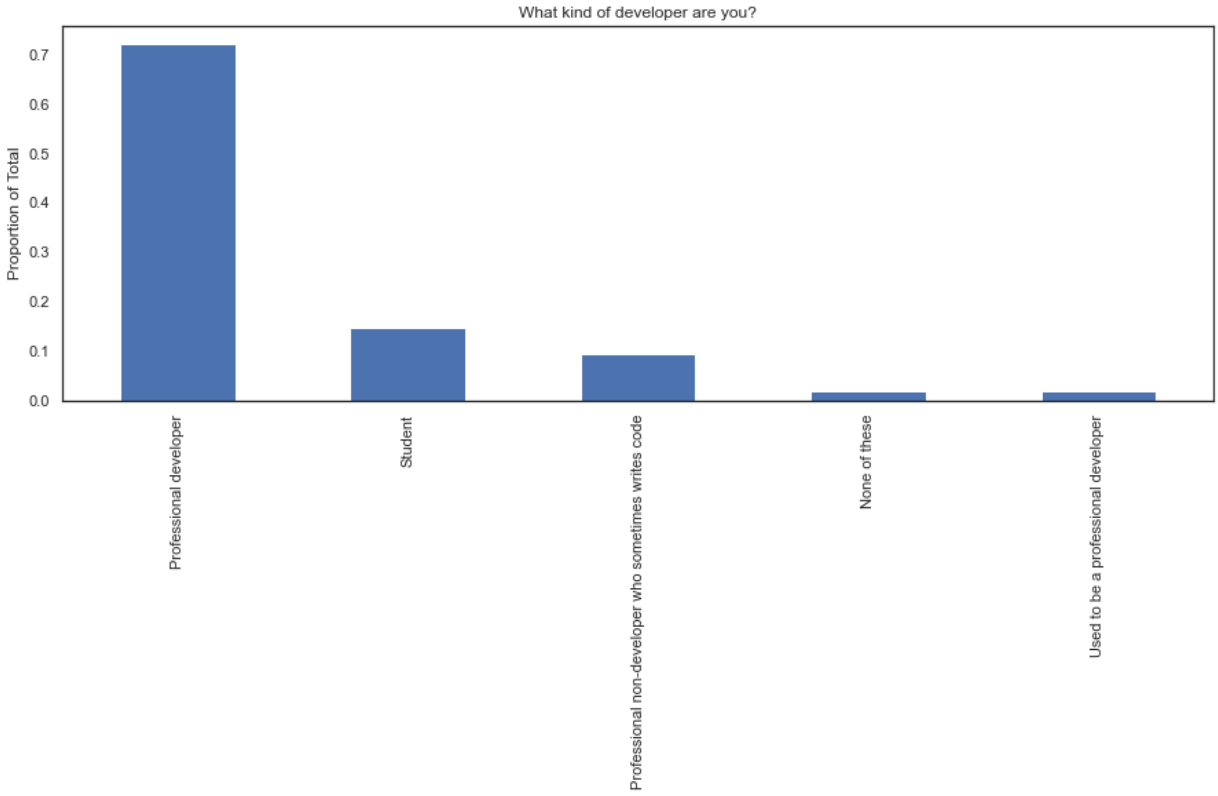
```
# Plot filtered counts by job satisfactoin and Job Seeking status broken by Gender a
sns.set(style="white", font_scale = 1)

chart = sns.relplot(x='JobSatisfaction',
                    y='JobSeekingStatus',
                    hue='WorldRegion',
                    size='Respondent',
                    sizes=(0, 3000),
                    alpha=.5,
                    #palette="muted",
                    height=5,
                    aspect = 1.5,
                    legend = 'brief',
                    col = 'Gender',
                    row = 'WorldRegion',
                    size_order = ('EasternEurope', 'Europe', 'North America', 'South
                    data=count)
#chart.set_title('Job Seeking Status vs. Job Satisfaction by Region and Gender')
```



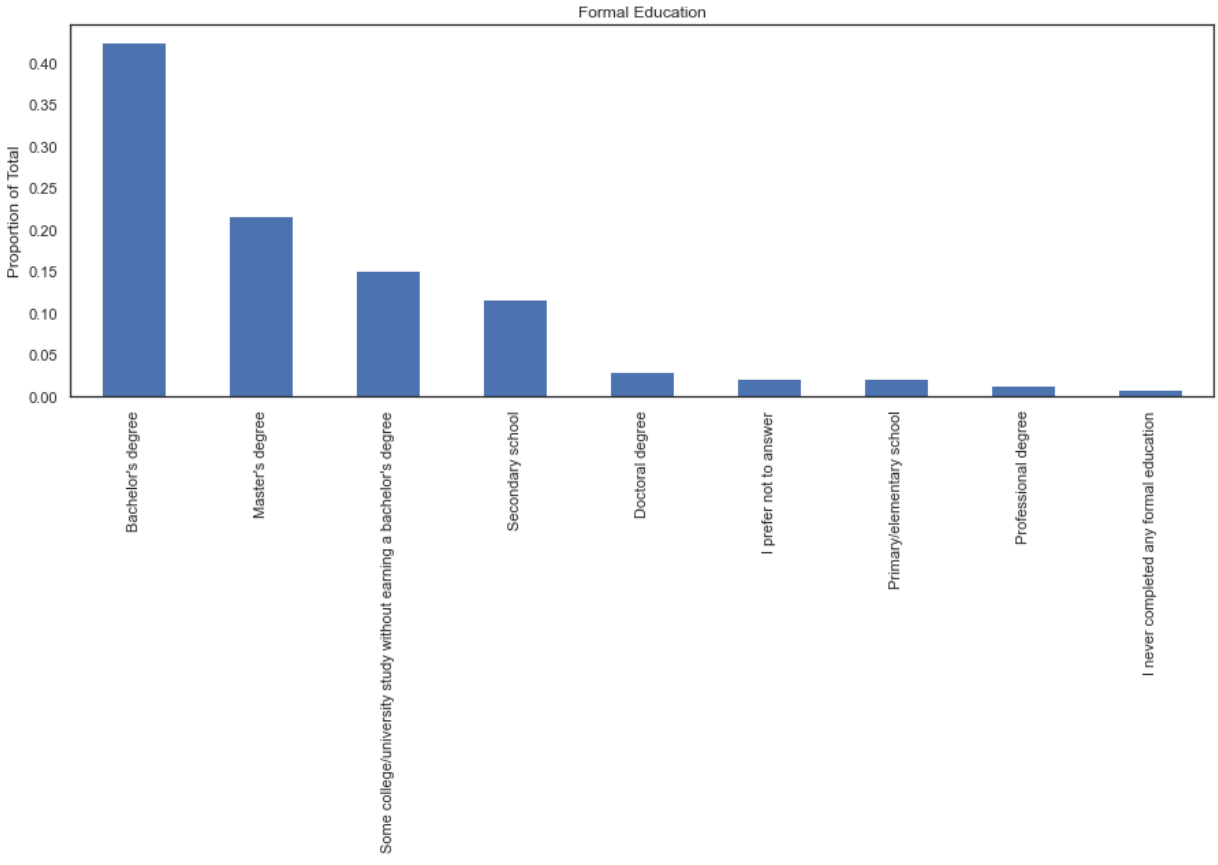

The Above 4 charts show the distribution of the job Seeking Status and job Satisfaction at survey level and the correlation between the 2. The last chart above does provide a Gender Break down as well and will be used lateron in this project.

The following charts show what kind of jobs and formal education the respondents have.

In [23]:
```
#Provide a pandas series of the counts for each Professional status

status_vals = df.Professional.value_counts()#Provide a pandas series of the counts f

plt.figure(figsize=(15,5))
(status_vals/df.shape[0]).plot(kind="bar")
plt.ylabel("Proportion of Total")
plt.title("What kind of developer are you?");
```




In [24]:
```
# counts for each FormalEducation status
ed_vals = df.FormalEducation.value_counts()

# a bar chart of the proportion of individuals' Formal Education
plt.figure(figsize=(15,5))
(ed_vals/df.shape[0]).plot(kind="bar")
plt.ylabel("Proportion of Total")
plt.title("Formal Education");
```




Digging in deeper looking at professionals and the 4 world regions of interest:

In [25]:
```
# Among all professionals, developers or non-developers

# Find Number of Respondents by Region
regionalRespondents = pd.DataFrame(df.query("(Professional == 'Professional develope
                                             Professional == 'Professional non-developer
                                             and \
                                            (WorldRegion == 'Europe' or \
                                             WorldRegion == 'North America' or \
                                             WorldRegion == 'Eastern Europe' or \
                                             WorldRegion == 'South Central Asia')"). \
                         groupby(['WorldRegion']).count()['Professional'])

# Determine job seeking status
regionalJobSeekers = pd.DataFrame(df.query("(Professional == 'Professional developer
                                            Professional == 'Professional non-developer who
                                            and \
                                           (WorldRegion == 'Europe' or \
                                            WorldRegion == 'North America' or \
                                            WorldRegion == 'Eastern Europe' or \
                                            WorldRegion == 'South Central Asia')"). \
                        groupby(['WorldRegion', 'JobSeekingStatus']).count()['Responden

# Display Job Satisfaction based on Job Seeking Status by world region
regionalSatisfaction = regionalJobSeekers.join(regionalRespondents)
regionalSatisfaction['regionalRatio'] = 100 * regionalSatisfaction['Respondent'] / r
regionalSatisfaction
```

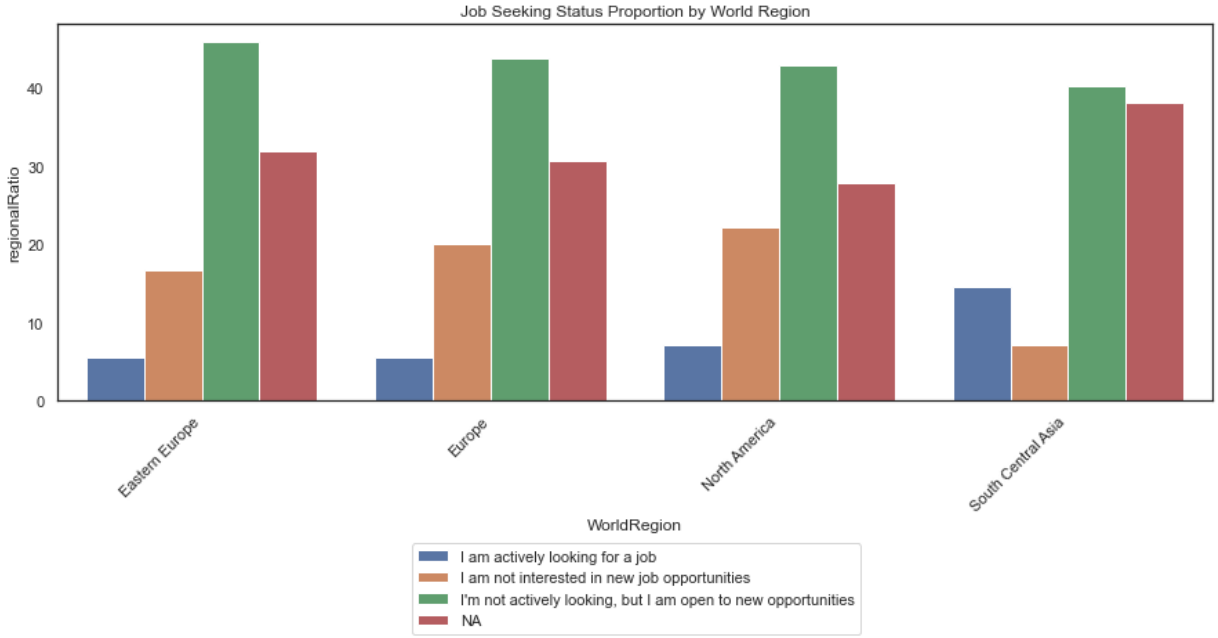
Out[25]:

| WorldRegion | JobSeekingStatus | Respondent | Professional | regionalRatio |
|---|---|---|---|---|
| Eastern Europe | I am actively looking for a job | 95 | 1697 | 5.598114 |
| | I am not interested in new job opportunities | 283 | 1697 | 16.676488 |
| | I'm not actively looking, but I am open to new opportunities | 778 | 1697 | 45.845610 |
| | NA | 541 | 1697 | 31.879788 |
| Europe | I am actively looking for a job | 285 | 5224 | 5.455590 |
| | I am not interested in new job opportunities | 1050 | 5224 | 20.099541 |
| | I'm not actively looking, but I am open to new opportunities | 2283 | 5224 | 43.702144 |
| | NA | 1606 | 5224 | 30.742726 |
| North America | I am actively looking for a job | 310 | 4385 | 7.069555 |
| | I am not interested in new job opportunities | 976 | 4385 | 22.257697 |
| | I'm not actively looking, but I am open to new opportunities | 1879 | 4385 | 42.850627 |
| | NA | 1220 | 4385 | 27.822121 |
| South Central Asia | I am actively looking for a job | 212 | 1458 | 14.540466 |
| | I am not interested in new job opportunities | 103 | 1458 | 7.064472 |

| | | Respondent | Professional | regionalRatio |
|---|---|---|---|---|
| **WorldRegion** | **JobSeekingStatus** | | | |
| | **I'm not actively looking, but I am open to new opportunities** | 587 | 1458 | 40.260631 |
| | **NA** | 556 | 1458 | 38.134431 |

In [26]:
```
# Resetting Index
regionalSatisfaction.reset_index(inplace=True)
```

In [27]:
```
plt.figure(figsize=(15,5))
chart = sns.barplot(data = regionalSatisfaction,
            x = 'WorldRegion',
            y = 'regionalRatio',
            hue = 'JobSeekingStatus',
            ci = None)
chart.set_title('Job Seeking Status Proportion by World Region')
chart.set_xticklabels(chart.get_xticklabels(), rotation=45, horizontalalignment='rig
chart.legend(loc='center left', bbox_to_anchor=(0.3, -0.5), ncol=1);
```




In [28]:
```
# Average Job Satisfaction by World Region

JSWR = pd.DataFrame(df.query("(Professional == 'Professional developer' or \
                               Professional == 'Professional non-developer who
                               and \
                              (WorldRegion == 'Europe' or \
                               WorldRegion == 'North America' or \
                               WorldRegion == 'Eastern Europe' or \
                               WorldRegion == 'South Central Asia')"). \
                    groupby(['WorldRegion']).mean()['JobSatisfaction'])


JSWR
```

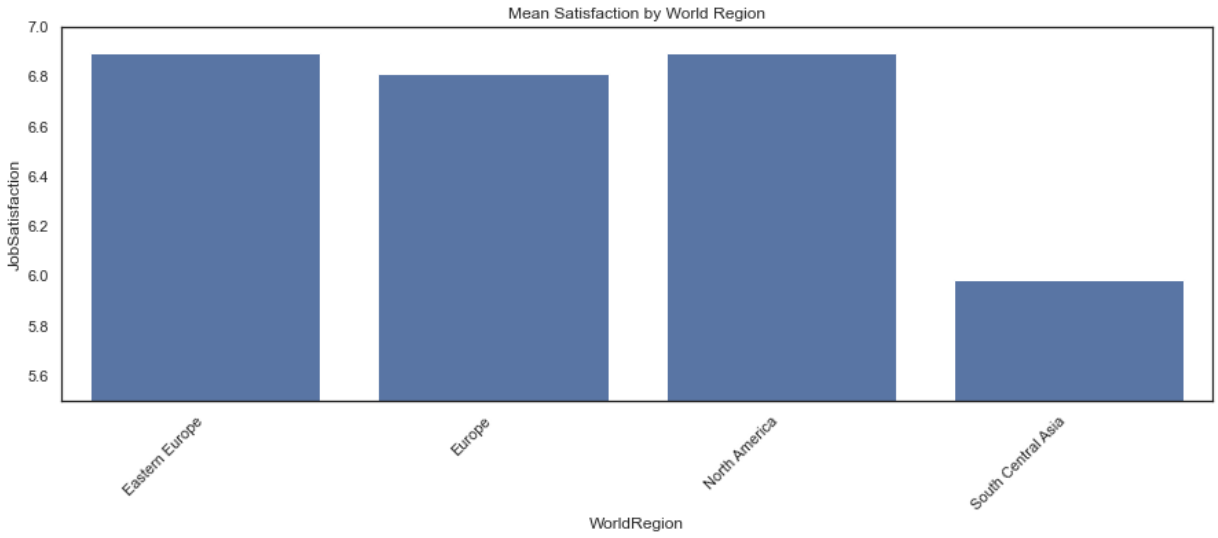
Out[28]:

| | JobSatisfaction |
|---|---|
| **WorldRegion** | |
| **Eastern Europe** | 6.896288 |
| **Europe** | 6.813744 |

| WorldRegion | JobSatisfaction |
|---|---|
| **North America** | 6.893957 |
| **South Central Asia** | 5.980796 |

In [29]:
```
# resetting index
JSWR.reset_index(level=0, inplace=True)
```

In [30]:
```
plt.figure(figsize=(15,5))
chart = sns.barplot(data = JSWR,
            x = 'WorldRegion',
            y = 'JobSatisfaction',
            #hue = 'JobSeekingStatus',
            ci = None,
            color = 'b')
chart.set_title('Mean Satisfaction by World Region')
chart.set_xticklabels(chart.get_xticklabels(), rotation=45, horizontalalignment='rig
chart.set(ylim=(5.5, 7));
#chart.legend(loc='center left', bbox_to_anchor=(0.3, -0.5), ncol=1);
```




In [31]:
```
# Average Job Satisfaction by World Region and Job Seeking Status

JSWR_JSS = pd.DataFrame(df.query("(Professional == 'Professional developer' or \
                                   Professional == 'Professional non-developer who
                                   and \
                                  (WorldRegion == 'Europe' or \
                                   WorldRegion == 'North America' or \
                                   WorldRegion == 'Eastern Europe' or \
                                   WorldRegion == 'South Central Asia')"). \
                        groupby(['WorldRegion', 'JobSeekingStatus']).mean()['JobSatisfa

JSWR_JSS
```

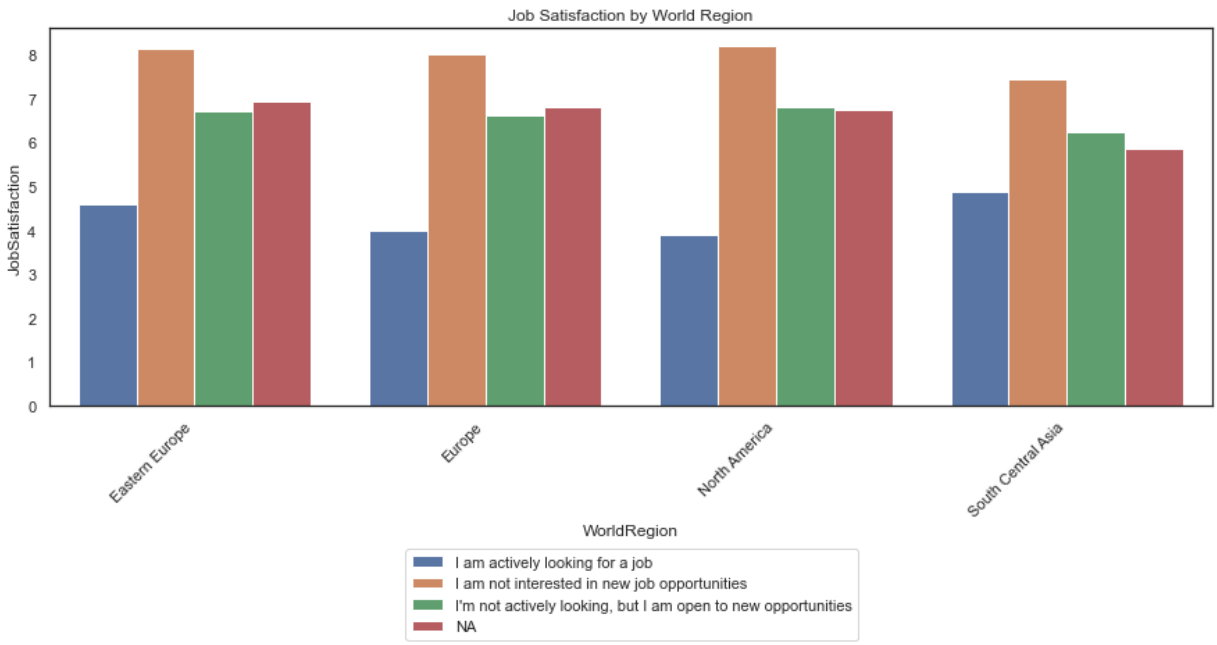
Out[31]:

| WorldRegion | JobSeekingStatus | JobSatisfaction |
|---|---|---|
| **Eastern Europe** | **I am actively looking for a job** | 4.589474 |
| | **I am not interested in new job opportunities** | 8.141343 |
| | **I'm not actively looking, but I am open to new opportunities** | 6.704370 |

| | | JobSatisfaction |
|---|---|---|
| **WorldRegion** | **JobSeekingStatus** | |
| | **NA** | 6.926063 |
| **Europe** | **I am actively looking for a job** | 3.989474 |
| | **I am not interested in new job opportunities** | 7.999048 |
| | **I'm not actively looking, but I am open to new opportunities** | 6.618046 |
| | **NA** | 6.818182 |
| **North America** | **I am actively looking for a job** | 3.903226 |
| | **I am not interested in new job opportunities** | 8.187500 |
| | **I'm not actively looking, but I am open to new opportunities** | 6.813199 |
| | **NA** | 6.743443 |
| **South Central Asia** | **I am actively looking for a job** | 4.886792 |
| | **I am not interested in new job opportunities** | 7.446602 |
| | **I'm not actively looking, but I am open to new opportunities** | 6.235094 |
| | **NA** | 5.857914 |

In [32]:
```
# Resetting Index
JSWR_JSS.reset_index(inplace=True)
```

In [33]:
```
plt.figure(figsize=(15,5))
chart = sns.barplot(data = JSWR_JSS,
            x = 'WorldRegion',
            y = 'JobSatisfaction',
            hue = 'JobSeekingStatus',
            ci = None)
chart.set_title('Job Satisfaction by World Region')
chart.set_xticklabels(chart.get_xticklabels(), rotation=45, horizontalalignment='rig
chart.legend(loc='center left', bbox_to_anchor=(0.3, -0.5), ncol=1);
```



From the data we can see that the best regions to be in as a professional are Eastern Europe, Europe and North America. The number of people looking for a job is the lowest and those who are not interested in a new job the highest, all while the number of people open to new opportunities is high.

South Central Asia and Eastern Europe do seem to have less tolerance to bad work conditions as among people who are looking for a job they have the highest job satisfaction.

Comparing Europe and North America is a little bit more difficult. While Job satisfaction among all catgeories from job seeking status are similar between the 2 regions, Europe does have a lower number of people looking for a job. North America, in contrast, has a higher number of people not looking for a job and both regions compare in looking out for new opportunities.

If we were to add the proportion of people not looking for a job and the non respondents for each of these 2 regions, which could be assumed becasue the the job satisfaction fro those non respondents is quite high for both regions, Europe would come out with a slight edge.

In conclusion, if somebody is trying to decide to chose the best world region to work in, Eastern Europe does seem to be an attractive option. If that is not aconsideration, it is pretty much a coin toss between Europe and North America.

In the end it may just come down to lifestyle and culture.

# MODULE 2: GENDER INFLUENCE ON JOB SATISFACTION

In [34]:
```
# Among all professionals, developers or non-developers and Males and Females

# Find Number of Respondents by Region and Gender
regionalRespondents_G = pd.DataFrame(df.query("(Professional == 'Professional develo
                                               and \
                                               (Gender == 'Male' or Gender == 'Female')\
                                               and\
                                              (WorldRegion == 'Europe' or WorldRegion == 'N
                                               WorldRegion == 'Eastern Europe' or WorldRegi
                        groupby(['WorldRegion', 'Gender']).count()['Professional'])

# Determine job seeking status
regionalJobSeekers_G = pd.DataFrame(df.query("(Professional == 'Professional develop
                                              and \
                                              (Gender == 'Male' or Gender == 'Female')\
                                              and\
                                             (WorldRegion == 'Europe' or WorldRegion == 'North
                                              WorldRegion == 'Eastern Europe' or WorldRegion =
                    groupby(['WorldRegion', 'JobSeekingStatus', 'Gender']).count()[


# Display Job Satisfaction based on Job Seeking Status by world region and Gender

#Joining the above 2 dataframes
regionalSatisfaction_G = regionalJobSeekers_G.join(regionalRespondents_G)

'''
Adding column representing percentage of professionals  per:
    - World region
```

```
    - Gender
    - Job Seeking Status
at World Region and Gender level
'''

regionalSatisfaction_G['regionalRatio_G'] = 100 * regionalSatisfaction_G['Respondent
regionalSatisfaction_G
```

Out[34]:

| | | | Respondent | Professional | regionalRatio_G |
|---|---|---|---|---|---|
| **WorldRegion** | **Gender** | **JobSeekingStatus** | | | |
| **Eastern Europe** | **Female** | **I am actively looking for a job** | 5 | 87 | 5.747126 |
| | | **I am not interested in new job opportunities** | 19 | 87 | 21.839080 |
| | | **I'm not actively looking, but I am open to new opportunities** | 43 | 87 | 49.425287 |
| | | **NA** | 20 | 87 | 22.988506 |
| | **Male** | **I am actively looking for a job** | 75 | 1106 | 6.781193 |
| | | **I am not interested in new job opportunities** | 207 | 1106 | 18.716094 |
| | | **I'm not actively looking, but I am open to new opportunities** | 558 | 1106 | 50.452080 |
| | | **NA** | 266 | 1106 | 24.050633 |
| **Europe** | **Female** | **I am actively looking for a job** | 17 | 238 | 7.142857 |
| | | **I am not interested in new job opportunities** | 68 | 238 | 28.571429 |
| | | **I'm not actively looking, but I am open to new opportunities** | 89 | 238 | 37.394958 |
| | | **NA** | 64 | 238 | 26.890756 |
| | **Male** | **I am actively looking for a job** | 197 | 3614 | 5.451024 |
| | | **I am not interested in new job opportunities** | 792 | 3614 | 21.914776 |
| | | **I'm not actively looking, but I am open to new opportunities** | 1741 | 3614 | 48.173769 |
| | | **NA** | 884 | 3614 | 24.460432 |
| **North America** | **Female** | **I am actively looking for a job** | 21 | 322 | 6.521739 |
| | | **I am not interested in new job opportunities** | 99 | 322 | 30.745342 |
| | | **I'm not actively looking, but I am open to new opportunities** | 156 | 322 | 48.447205 |
| | | **NA** | 46 | 322 | 14.285714 |
| | **Male** | **I am actively looking for a job** | 220 | 2946 | 7.467753 |
| | | **I am not interested in new job opportunities** | 698 | 2946 | 23.693143 |
| | | **I'm not actively looking, but I am open to new opportunities** | 1392 | 2946 | 47.250509 |
| | | **NA** | 636 | 2946 | 21.588595 |

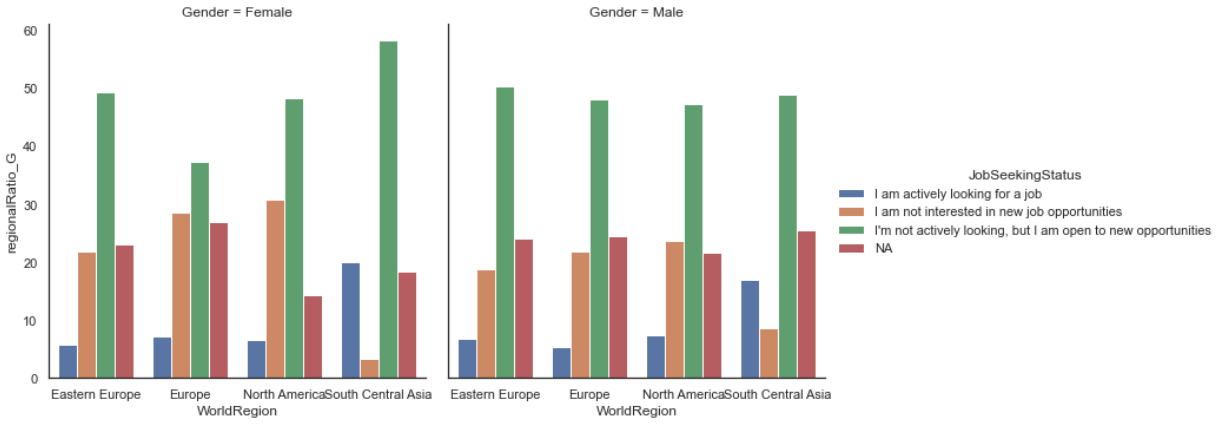
| | | | Respondent | Professional | regionalRatio_G |
|---|---|---|---|---|---|
| **WorldRegion** | **Gender** | **JobSeekingStatus** | | | |
| **South Central Asia** | **Female** | **I am actively looking for a job** | 12 | 60 | 20.000000 |
| | | **I am not interested in new job opportunities** | 2 | 60 | 3.333333 |
| | | **I'm not actively looking, but I am open to new opportunities** | 35 | 60 | 58.333333 |
| | | **NA** | 11 | 60 | 18.333333 |
| | **Male** | **I am actively looking for a job** | 119 | 701 | 16.975749 |
| | | **I am not interested in new job opportunities** | 60 | 701 | 8.559201 |
| | | **I'm not actively looking, but I am open to new opportunities** | 343 | 701 | 48.930100 |
| | | **NA** | 179 | 701 | 25.534950 |

Job Seeking Status by World Region and Gender:

In [35]:
```
#resetting index
regionalSatisfaction_G.reset_index(inplace=True)
```

In [36]:
```
plt.figure(figsize=(15,5))
chart = sns.catplot(data = regionalSatisfaction_G,
            x = 'WorldRegion',
            y = 'regionalRatio_G',
            hue = 'JobSeekingStatus',
            col = 'Gender',
            kind = 'bar',
            ci = None);
#chart.set_title('Job Seeking Status Proportion by World Region')
#chart.set_xticklabels(chart.get_xticklabels(), rotation=45, horizontalalignment='ri
#chart.legend(loc='center left', bbox_to_anchor=(0.3, -0.5), ncol=1);
```

```
<Figure size 1080x360 with 0 Axes>
```




Of all the regions of interest Women are happier with their current job than Men, with the exception of South Central Asia where this is switched around.

Also when looking at professional open to new opportunities Woemn in Europe areless interested and women in South Central Asia most interested, while in Eastern Europe and North America it is middle of the road.

While professionals actively looking for a job is between 5% and 7% across Europe and North America for both Genders, in SouthCentral Asia women are by far more active than other regions and over their male counterparts.

In [37]:
```
#  How does gender play a role?

#get average job satisfaction by world region and gender

JSWR_G = pd.DataFrame(df.query("(Professional == 'Professional developer' or Profess
JSWR_G
```

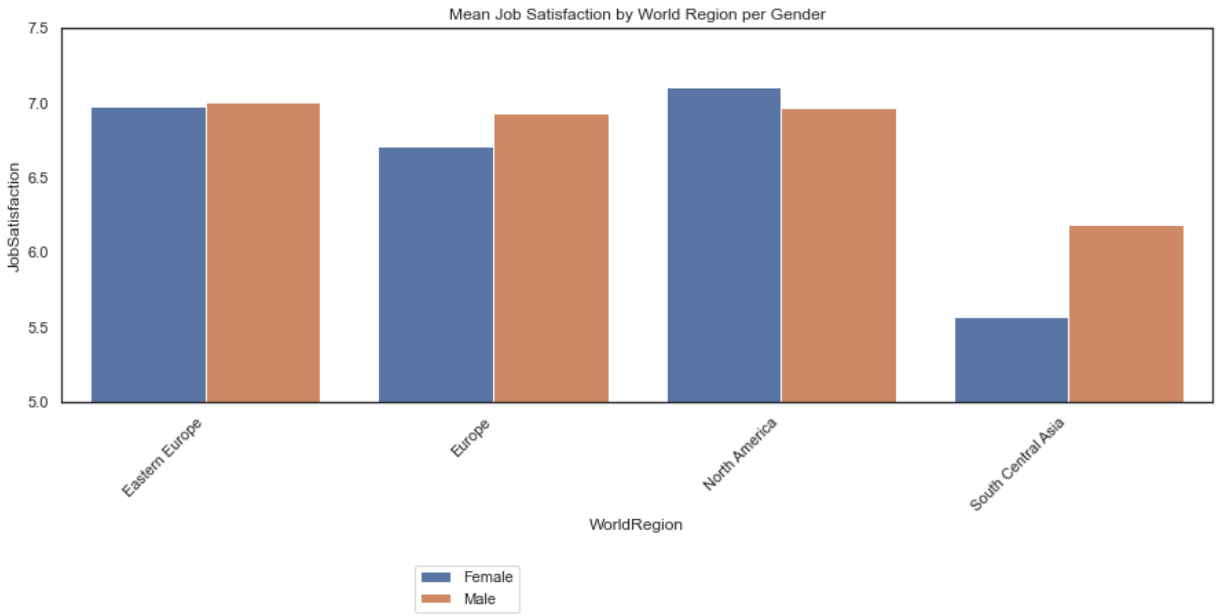
Out[37]:

| | | JobSatisfaction |
|---|---|---|
| **WorldRegion** | **Gender** | |
| **Eastern Europe** | **Female** | 6.977011 |
| | **Male** | 6.999096 |
| **Europe** | **Female** | 6.710084 |
| | **Male** | 6.932485 |
| **North America** | **Female** | 7.105590 |
| | **Male** | 6.967413 |
| **South Central Asia** | **Female** | 5.566667 |
| | **Male** | 6.188302 |

In [38]:
```
#resetting index
JSWR_G.reset_index(inplace=True)
```

Plotting Mean Job Satisfactoin by World Region and Gender:

In [39]:
```
plt.figure(figsize=(15,5))
chart = sns.barplot(data = JSWR_G,
            x = 'WorldRegion',
            y = 'JobSatisfaction',
            hue = 'Gender',
            ci = None)
chart.set_title('Mean Job Satisfaction by World Region per Gender')
chart.set_xticklabels(chart.get_xticklabels(), rotation=45, horizontalalignment='rig
chart.set(ylim=(5, 7.5));
chart.legend(loc='center left', bbox_to_anchor=(0.3, -0.5), ncol=1);
```




In [40]:
```
#get average job satisfaction by world region, job seeking status and gender

JSWR_G = pd.DataFrame(df.query("(Professional == 'Professional developer' or Profess

JSWR_G
```

Out[40]:

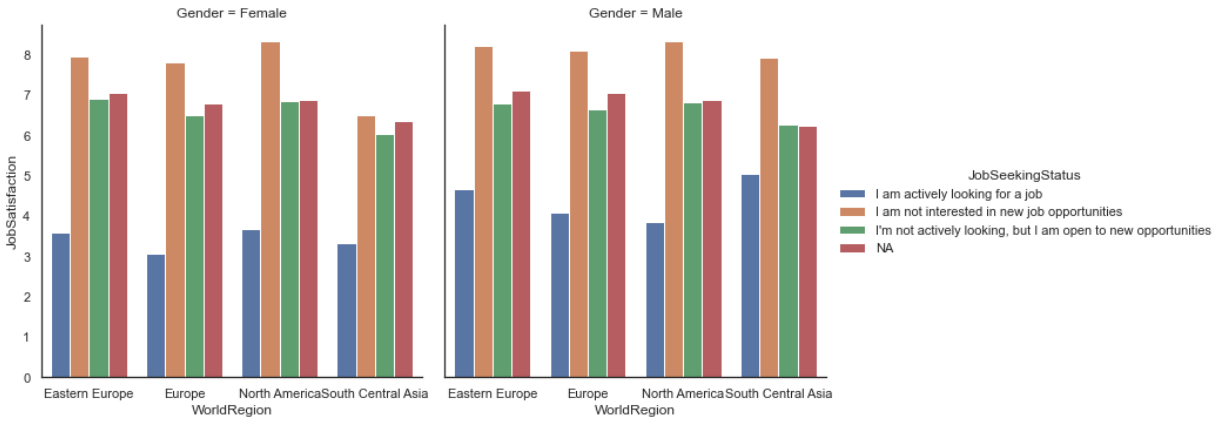
| | | | JobSatisfaction |
|---|---|---|---|
| **WorldRegion** | **JobSeekingStatus** | **Gender** | |
| **Eastern Europe** | **I am actively looking for a job** | **Female** | 3.600000 |
| | | **Male** | 4.680000 |
| | **I am not interested in new job opportunities** | **Female** | 7.947368 |
| | | **Male** | 8.227053 |
| | **I'm not actively looking, but I am open to new opportunities** | **Female** | 6.906977 |
| | | **Male** | 6.802867 |
| | **NA** | **Female** | 7.050000 |
| | | **Male** | 7.109023 |
| **Europe** | **I am actively looking for a job** | **Female** | 3.058824 |
| | | **Male** | 4.076142 |
| | **I am not interested in new job opportunities** | **Female** | 7.808824 |
| | | **Male** | 8.104798 |
| | **I'm not actively looking, but I am open to new opportunities** | **Female** | 6.505618 |
| | | **Male** | 6.659966 |
| | **NA** | **Female** | 6.796875 |
| | | **Male** | 7.055430 |
| **North America** | **I am actively looking for a job** | **Female** | 3.666667 |
| | | **Male** | 3.854545 |
| | **I am not interested in new job opportunities** | **Female** | 8.323232 |
| | | **Male** | 8.335244 |

| | | | JobSatisfaction |
|---|---|---|---|
| **WorldRegion** | **JobSeekingStatus** | **Gender** | |
| | **I'm not actively looking, but I am open to new opportunities** | **Female** | 6.865385 |
| | | **Male** | 6.812500 |
| | **NA** | **Female** | 6.869565 |
| | | **Male** | 6.882075 |
| **South Central Asia** | **I am actively looking for a job** | **Female** | 3.333333 |
| | | **Male** | 5.033613 |
| | **I am not interested in new job opportunities** | **Female** | 6.500000 |
| | | **Male** | 7.933333 |
| | **I'm not actively looking, but I am open to new opportunities** | **Female** | 6.028571 |
| | | **Male** | 6.256560 |
| | **NA** | **Female** | 6.363636 |
| | | **Male** | 6.240223 |

In [41]:
```
# resetting indexes
JSWR_G.reset_index(inplace=True)
```

In [42]:
```
plt.figure(figsize=(15,5))
chart = sns.catplot(data = JSWR_G,
                    x = 'WorldRegion',
                    y = 'JobSatisfaction',
                    hue = 'JobSeekingStatus',
                    col = 'Gender',
                    kind = 'bar',
                    ci = None);
#chart.set_title('Job Seeking Status Proportion by World Region')
#chart.set_xticklabels(chart.get_xticklabels(), rotation=45, horizontalalignment='ri
#chart.legend(loc='center left', bbox_to_anchor=(0.3, -0.5), ncol=1);
```

```
<Figure size 1080x360 with 0 Axes>
```




In [43]:
```
# resetting indexes
regionalSatisfaction.reset_index(inplace=True)
```

At a high level job satisfaction for women is lower than their male counterparts, except for North America where women seem to be happier with their jobs. In South Central Asia women are dissatisfied to a hihger degree than their counterparts in other regions.

Women across the board need to be more dissatisfied with their jobs before starting to look for a job, but are just as open to new opportunities as men.

Women in Europe on the other hand, while they seem more dissatisfied with their jobs, are less interested in moving around, as opposed to their eastern european colleagues, while more satisfied are also more mobile.

For Women North America is the best place to be. Higher job satisfaction, more opportunites and less desire to move around are indications of a better professional life for women in this region.

# MODULE 3: ROLE OF FORMAL EDUCATION IN JOB SATISFACTION

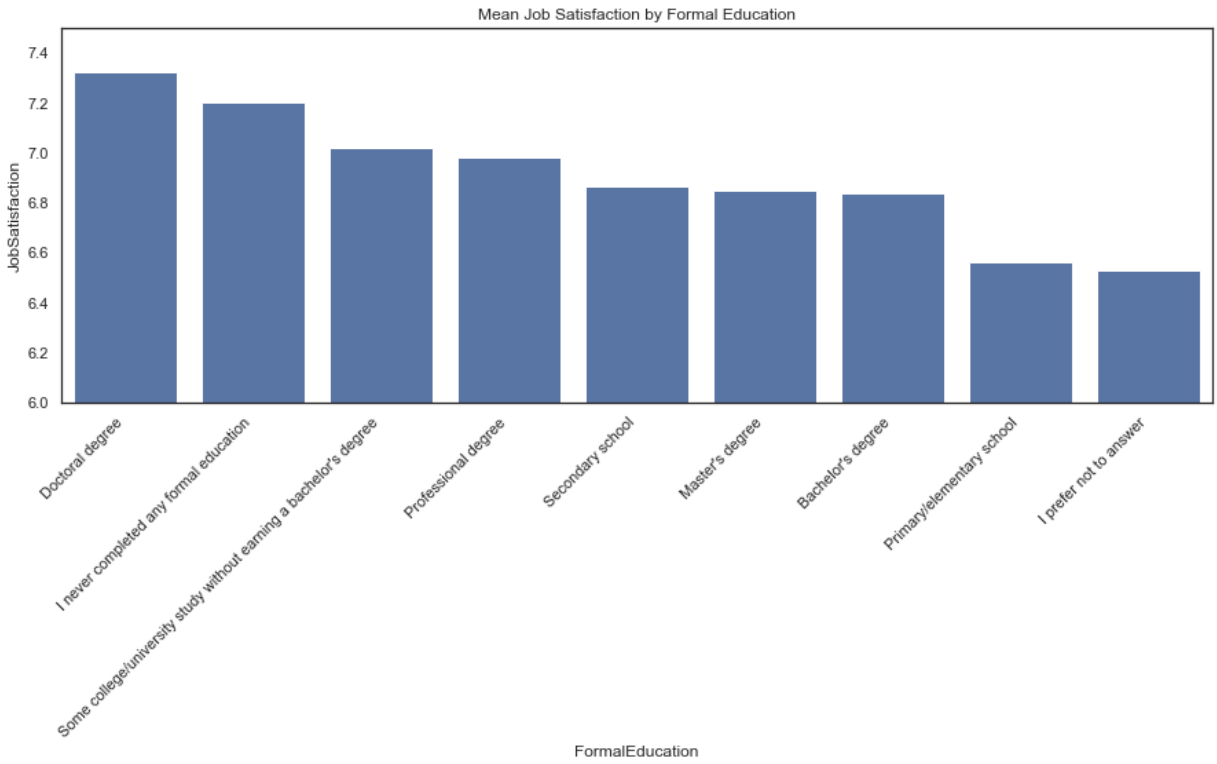
In [44]:
```
#get average job satisfaction by Formal Education for the survey

JSE = pd.DataFrame(df.query("(Professional == 'Professional developer' or Profession
JSE_sorted = JSE.sort_values('JobSatisfaction', ascending=False)
JSE_sorted
```

Out[44]:

| FormalEducation | JobSatisfaction |
|---|---|
| **Doctoral degree** | 7.322581 |
| **I never completed any formal education** | 7.200000 |
| **Some college/university study without earning a bachelor's degree** | 7.021355 |
| **Professional degree** | 6.980392 |
| **Secondary school** | 6.864686 |
| **Master's degree** | 6.845161 |
| **Bachelor's degree** | 6.835884 |
| **Primary/elementary school** | 6.560976 |
| **I prefer not to answer** | 6.525424 |

In [45]:
```
# Resetting INdex
JSE_sorted.reset_index(level=0, inplace=True)
```

In [46]:
```
# plot Job Satisfaction by formal education for the survey

plt.figure(figsize=(15,5))
chart = sns.barplot(data = JSE_sorted,
            x = 'FormalEducation',
            y = 'JobSatisfaction',
            ci = None,
            color = 'b')
chart.set_title('Mean Job Satisfaction by Formal Education')
chart.set_xticklabels(chart.get_xticklabels(), rotation=45, horizontalalignment='rig
chart.set(ylim=(6, 7.5));
#chart.legend(loc='center left', bbox_to_anchor=(0.3, -0.5), ncol=1);
```




In [47]:
```
#get average job satisfaction by world region and Formal Education

JS_edu = pd.DataFrame(df.query("(Professional == 'Professional developer' or Profess
JS_edu
JS_edu.reset_index(inplace=True)

JS_edu
```

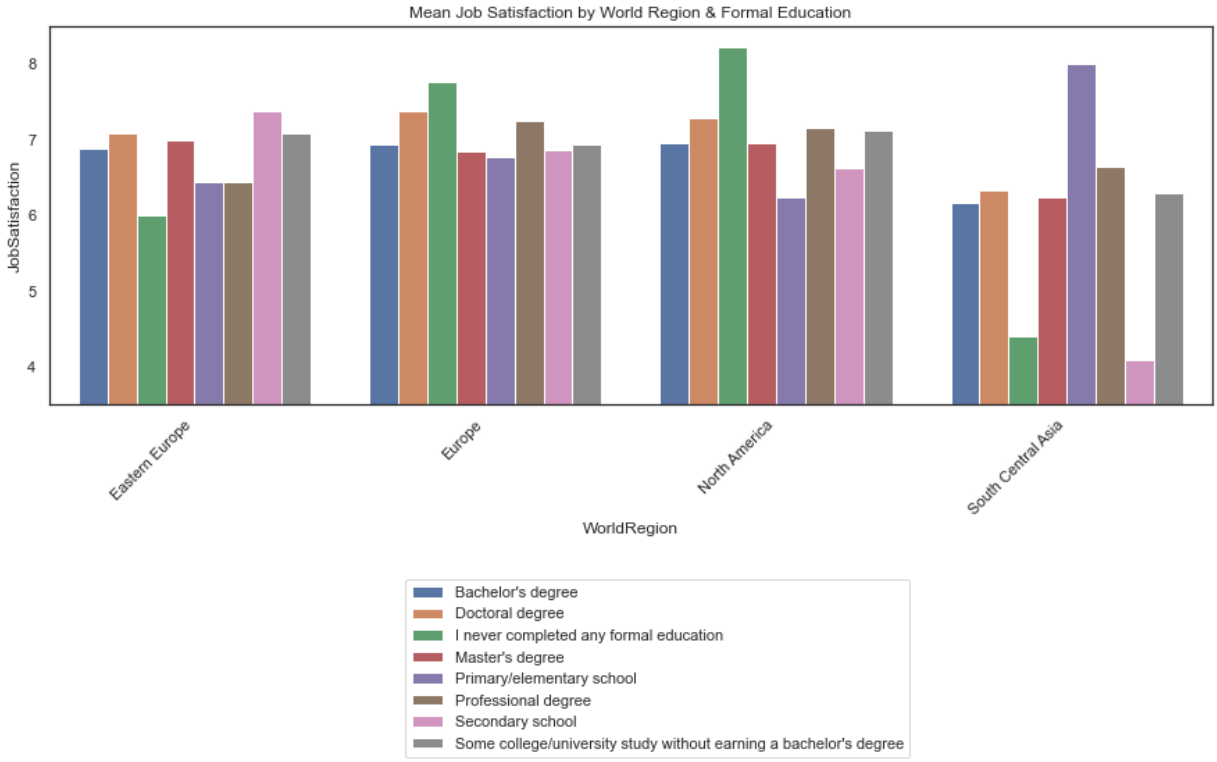
Out[47]:

| | WorldRegion | FormalEducation | JobSatisfaction |
|---|---|---|---|
| **0** | Eastern Europe | Bachelor's degree | 6.879607 |
| **1** | Eastern Europe | Doctoral degree | 7.086957 |
| **2** | Eastern Europe | I never completed any formal education | 6.000000 |
| **3** | Eastern Europe | Master's degree | 6.988345 |
| **4** | Eastern Europe | Primary/elementary school | 6.444444 |
| **5** | Eastern Europe | Professional degree | 6.434783 |
| **6** | Eastern Europe | Secondary school | 7.369748 |
| **7** | Eastern Europe | Some college/university study without earning ... | 7.085366 |
| **8** | Europe | Bachelor's degree | 6.940665 |
| **9** | Europe | Doctoral degree | 7.379808 |
| **10** | Europe | I never completed any formal education | 7.764706 |
| **11** | Europe | Master's degree | 6.834842 |
| **12** | Europe | Primary/elementary school | 6.777778 |
| **13** | Europe | Professional degree | 7.250000 |
| **14** | Europe | Secondary school | 6.864641 |
| **15** | Europe | Some college/university study without earning ... | 6.937173 |
| **16** | North America | Bachelor's degree | 6.948108 |

| | WorldRegion | FormalEducation | JobSatisfaction |
|---|---|---|---|
| **17** | North America | Doctoral degree | 7.289720 |
| **18** | North America | I never completed any formal education | 8.222222 |
| **19** | North America | Master's degree | 6.945946 |
| **20** | North America | Primary/elementary school | 6.230769 |
| **21** | North America | Professional degree | 7.150000 |
| **22** | North America | Secondary school | 6.628319 |
| **23** | North America | Some college/university study without earning ... | 7.117845 |
| **24** | South Central Asia | Bachelor's degree | 6.161959 |
| **25** | South Central Asia | Doctoral degree | 6.333333 |
| **26** | South Central Asia | I never completed any formal education | 4.400000 |
| **27** | South Central Asia | Master's degree | 6.235294 |
| **28** | South Central Asia | Primary/elementary school | 8.000000 |
| **29** | South Central Asia | Professional degree | 6.636364 |
| **30** | South Central Asia | Secondary school | 4.083333 |
| **31** | South Central Asia | Some college/university study without earning ... | 6.296296 |

In [48]:

```
# produce a plot  shoing job satisfaction by World region and formal education

plt.figure(figsize=(15,5))
sns.set(style="white", font_scale = 1)
chart = sns.barplot(data = JS_edu,
            x = 'WorldRegion',
            y = 'JobSatisfaction',
            hue = 'FormalEducation',
            #height = 8,
            #aspect = 3,
            ci = None
            )
chart.set_title('Mean Job Satisfaction by World Region & Formal Education')
chart.set_xticklabels(chart.get_xticklabels(), rotation=45, horizontalalignment='rig
chart.legend(loc='center left', bbox_to_anchor=(0.3, -0.7), ncol=1)
chart.set(ylim=(3.5, 8.5));
```




Overall Job Satisfaction is highest among people with a Doctoral degree, followed by people with no formal education or people with some college but no degree.

When taking into account the world region this seems to hold true for North America and Europe. Professionals in these 2 regions have good chances at a good job even if they have no formal college degree.

Europe has better job satisfaction across education levels, indicating there are good jobs for professionals form all walks of life. In North America howeever, professionals with only primary or high school level educatoin are not as happy.

Europe is the place to be for professionals at all levels of education. In North America at least some college or a professional degree would be required to have a satisfying job although those without completing a fomral education are the most satisfied.

Eastern Europe is also a region to conisder, South Central Asia however has the least satsified people across the board , except for those with only primary school education, who have the highest satisfaction across all regions.

---

**J - Component Project for MGT1025 at VIT, Vellore**