

НИС №10 “Работа по проекту анализа данных IoT в группах”

```
def normalization():
    data = pd.read_csv('clean_holes.csv', sep=';', engine='python')
    df = pd.DataFrame(data)

    print(f"DataFrame:\n{df}\n")
    print(f"column types:\n{df.dtypes}")

    holes_list = []

    col_List= df['Скважина'].tolist()
    num = 1
    for i in range(len(col_List)):
        if (i == 0) or (col_List[i] != col_List[i-1]):

            holes_list.append(col_List[i])
            num += 1

    columns = ['Рпн(ТМ)', 'Рзаб(Рпн)', 'Рзаб(Нд)', 'Рзаб(иссл)']

    for hole in holes_list:
        df1 = df[lambda df: df['Скважина'] == hole]

        for col in columns:
            col_List= df1[col].tolist()

            for i in range(len(col_List)):
                delta = len(col_List) - i
                left = col_List[i-1]+col_List[i-2]+col_List[i-3]+col_List[i-4]+col_List[i-5]
                if delta == 1:
                    right = col_List[0]+col_List[1]+col_List[2]+col_List[3]+col_List[4]
                if delta == 2:
                    right = col_List[i+1]+col_List[0]+col_List[1]+col_List[2]+col_List[3]
                if delta == 3:
                    right = col_List[i+1]+col_List[i+2]+col_List[0]+col_List[1]+col_List[2]
                if delta == 4:
                    right = col_List[i+1]+col_List[i+2]+col_List[i+3]+col_List[0]+col_List[1]
                if delta == 5:
                    right = col_List[i+1]+col_List[i+2]+col_List[i+3]+col_List[i+4]+col_List[0]
                if delta >= 6:
                    right = col_List[i+1]+col_List[i+2]+col_List[i+3]+col_List[i+4]+col_List[i+5]
                col_List[i] = (left + right) / 10

            df1[col] = col_List
            df[lambda df: df['Скважина'] == hole] = df1

    max = df.max()
    df['Рпн(ТМ)'] = df['Рпн(ТМ)'].apply(lambda x: round(x / max['Рпн(ТМ)'], 3))
```

```
df['Рззб(Рнр)'] = df['Рззб(Рнр)'].apply(lambda x: round(x / max['Рззб(Рнр)'], 3))
df['Рззб(Нд)'] = df['Рззб(Нд)'].apply(lambda x: round(x / max['Рззб(Нд)'], 3))
df['Рззб(иссл)'] = df['Рззб(иссл)'].apply(lambda x: round(x / max['Рззб(иссл)'], 3))

df.to_csv('normal_holes.csv', index=False, sep=';', encoding='cp1251')
```

