

**Problem 3 (20 points).** Soundex is an algorithm for coding names. It was developed in the USA in 1918–1922 by Robert C. Russell and Margaret King Odell in order to facilitate searching for similar-sounding surnames. In the middle of the 20th century, Soundex was extensively used in the USA to analyze results of 1890–1920 censuses.

Below is a sample card with data from the 1910 census. You can see the Soundex code for *Wilson*, which is W425.

HEAD OF FAMILY			E.D.	LOUISIANA SHEET
<i>W 425 Wilson, Alice</i>			118	17
COLOR	AGE	BIRTHPLACE		
B	42			
COUNTY	St. Landry		CITY	
OTHER MEMBERS OF FAMILY				
NAME	RELATIONSHIP	AGE	BIRTHPLACE	
<i>Eugene</i>	<i>W</i>	<i>46</i>		
<i>Regina</i>	<i>R</i>	<i>15</i>		
<i>Walter</i>	<i>W</i>	<i>13</i>		
<i>Louise</i>	<i>R</i>	<i>12</i>		
<i>Carroll</i>	<i>R</i>	<i>7</i>		
<i>Capell</i>	<i>S</i>	<i>7</i>		
<i>Hudson</i>	<i>S</i>	<i>4</i>		
FORM 10-636 (4-20-61) 1910 CENSUS INDEX - FAMILY				
U.S. DEPARTMENT OF COMMERCE BUREAU OF THE CENSUS				

Source: <https://familysearch.org/learn/wiki/en/Soundex>

Here is a list of surnames, with the corresponding Soundex codes in arbitrary order. Some characters are missing:

*Allaway, Anderson, Ashcombe, Buckingham, Chapman, Colquhoun, Evans, Fairwright, Kingscott, Lewis, Littlejohns, Stanmore, Stubbs, Tocher, Tonks, Whytehead*      S312, T6u, 53, C42u, T520, L42, A536, C155, 623, S356, 252, 152, 330, A251, A400, L20

- (a) Describe how a Soundex code is produced, step by step.
- (b) Match the surnames with the corresponding Soundex codes and restore the omitted characters.
- (c) Generate Soundex codes for the following surnames:

*Ferguson, Fitzgerald, Hamnett, Keefe, Maxwell, Razey, Shaw, Upfield.*

—Alexander Piperski