* Soundex augosithm soundex is considered as a Phonetic agosithm used Primusity in NLP 508 indexing names by sound -) Algorithm used to group similar sounding letters together and assign each group a numeral number -) main good of this technique is to use homophones for encoding text with numerical indexing - As a sessut the numerical separentation can easily matched with other similar sounding characters having same numesiccu code - This sesults in setsiving a list of words that are pronounced similarly with very little variation in their homophones - Used as a simple phonetic based spell Checkes -) code sol a word consists of its siest letter followed by three numbers that encode the semuining consonant -) Those consonants that generate sume sound have same number

-) uses code to check for closest work

mul to Letters digit substitute with integes B, F, P, V C, K, J, K, S, X, Z O, T MIN Algositam - Remove all punctuation masks & capitalize letters in the word -) Retain the sisst letter of the word - Remove any occusence of the letters - A, E, I, O, U, H, W, Y UPart from very first letter -) Replace the letters by numbers shown in tuble If two or more adjucent letters, not seperated by vowers, have the same numeric value, setain only one of them Also two letters with the same number seperated by 'h' or 'w' are codded as a single number where as such a letters seperated by a vower codded twice -) Return the first four characters, Rad with zero is there are less than four

PAGENO.

DATE: 1

	word sounder code
	V.9,7,2
	axate Cc630
	agent a630
	arate -> arate
	CCRT
	Plansing & marketing
	CC 63
2	Removasiable Beneral advance and
	150000 Dat CC &63001 DSUBLISH
Poor.	- Retain the easter letters not year
ectice	Execut -> CLREAT
2000	BOBE E DO DUHI WY REPART S
	GRTHSI 1200
mods	CC630 300 300 300 300 300 300 300 300 300

TORN -> TORN V TRN T65 March Dans T650 used to sind good Possible candidate WORN - 50650 for correction to be HORN -) H650 Effected fox a misspelled word # AShCRUST ASHCRAFT 4 two 08 more adjucent retters not seperated by vower have the same numeric vouve, retorn one of them A2613 I rerun the sixst sous characters A261



A	Tumczak Magan Son Magan
	TM CZK
	V. demonstration
	T5222 J setuin only one of them
	T520 020T
26 70	The Oth Social Conference of the State of th
A	Tiusnim oran araow
020	2 1 2 2 1 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
15	TJSNM
1201	
	T2255
	AS LONG COSTA DE COSTA DE SE LA SEL
	T250
	SON SAGNICE ENFORCEMENT OF SON
	and the state of t
	A DECRET COMMENTED
	The state of the s
	SWARD TELEVISION STORES OF THE
1500	DE THE SHOW TO OUR WAR VIEW
1831	
DANG	
10126	British Dissimila Burns 343
	ansat to suc
	The state of the s
	PROBLEM COC HARPY SOLON !